

# DECONSTRUCTING LANGUAGE STRUCTURE AND MEANING

*Studies on  
Syntax, Semantics  
and Phonology*

*Edited by*  
Mihaela Tănase-Dogaru  
Alina Tigău  
Mihaela Zamfirescu

# Deconstructing Language Structure and Meaning



# Deconstructing Language Structure and Meaning:

*Studies on Syntax, Semantics  
and Phonology*

Edited by

Mihaela Tănase-Dogaru, Alina Tigău  
and Mihaela Zamfirescu

**Cambridge  
Scholars  
Publishing**



Deconstructing Language Structure and Meaning:  
Studies on Syntax, Semantics and Phonology

Edited by Mihaela Tănase-Dogaru, Alina Tigău and Mihaela Zamfirescu

This book first published 2021

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data  
A catalogue record for this book is available from the British Library

Copyright © 2021 by Mihaela Tănase-Dogaru, Alina Tigău,  
Mihaela Zamfirescu and contributors

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-7001-0

ISBN (13): 978-1-5275-7001-6

# TABLE OF CONTENTS

Chapter One.....	1
Issues in Non-Finite Complementation: Control Jutta Hartmann (Bielefeld University)	
Chapter Two.....	26
Scope Dependencies with Romanian Ditransitives Alina Tigău (University of Bucharest)	
Chapter Three.....	45
The <i>i</i> * Single Argument Introducer: A Solution for Representing the Beneficiary Argument of Creation Verbs in Brazilian Portuguese Ana Regina Calindro (Federal University of Rio de Janeiro)	
Chapter Four.....	76
Datives in Istro-Romanian Ionuț Geană (Institutul de Lingvistică „Iorgu Iordan - Al. Rosetti”, București; Universitatea din Bucaresti Arizona State University)	
Chapter Five.....	92
(Bare) Objects of Prepositions in Eastern Romance Dana Isac (Concordia University)	
Chapter Six.....	114
The Finite Clause Boundedness of Quantifier Scope: Evidence from Hungarian Gergő Turi and Balázs Surányi (Pázmány Péter Catholic University; Research Institute for Linguistics — Hungarian Academy of Sciences)	
Chapter Seven.....	132
Polarity Sensitive Items in Non-veridical Contexts in Romanian Mihaela Zamfirescu (University of Bucharest)	
Chapter Eight.....	151
Two Types of Quantity Relative Superlatives Ion Giurgea (Institutul de Lingvistică „Iorgu Iordan - Al. Rosetti”)	

Chapter Nine.....	180
Romanian ‘câte’ as a Nominal and Adverbial Marker of Distributivity Mara Panaitescu (Bucharest University)	
Chapter Ten .....	198
Fortition in the Historical Phonology of Maltese: Two Case Studies Andrei A. Avram (University of Bucharest)	
Chapter Eleven .....	220
The English Positional Schwa: A Similar Category to Establish by Romanians Elena-Raluca Constantin (Military Technical Academy of Bucharest)	
Chapter Twelve .....	237
Some Notes on the Syntax of Corrective ‘ci’ / ‘but’ in Romanian Mihaela Tănase-Dogaru (University of Bucharest)	
Chapter Thirteen .....	257
Subordination and Fragments: The Case of Romanian Gabriela Bîlbîie (University of Bucharest)	
Chapter Fourteen .....	278
The Acquisition of Subjects in a Romanian-Hungarian Bilingual Context Veronica Tomescu (University of Bucharest)	
Chapter Fifteen .....	303
5-year-olds are Precise with Cardinals: Experimental Evidence from Romanian Child Language Camelia Bleotu (University of Bucharest)	
Chapter Sixteen .....	322
Experimental Data on Telicity Inferences in Romanian Ioana Stoicescu (Bucharest University)	
Contributors.....	342

# CHAPTER ONE

## ISSUES IN NON-FINITE COMPLEMENTATION: CONTROL

JUTTA M. HARTMANN

This paper argues for an analysis of control that integrates (lexical-) semantic and syntactic aspects. It shows that such an integration is necessary for cases as anti-subject control in German and the polysemous Dutch verb *zeggen*. Anti-subject-control is argued to require a lexical-semantic basis for an anti-control configuration, which correlates with a syntactic configuration that blocks structural control. For the Dutch verb *zeggen* which is polysemous when combined with infinitival complements, it is shown in a more detailed case study that the different readings correlate with different syntactic properties. A tentative analysis is provided, which takes these readings to be the result of a combination of a general verb SAY with different complement types.

Keywords: control, polysemy, infinitives, Dutch, German

### 1. Introduction

Languages vary in how they express propositional arguments in a broad sense (events, propositions, situations). In English, e.g., we find nominal or nominalized arguments (1), gerunds (2), infinitives of various types (3)-(5) as well as finite clauses (7).

- (1) a. Stella cannot afford to believe [the fact that Blanche could be right]. (BNC, HUB 386)
- b. Nigel enjoyed [her admiration of his writing]. (BNC, AC3 1162)



- (2) a. I resented [his constant questioning of my motives].  
[nominal]  
b. I resented [his constantly questioning my motives].  
[verbal]  
(Huddleston 2002b, 1189)
- (3) a. Liz hoped [to convince them]. [control]  
b. Liz seemed [to convince them]. [raising]  
(Huddleston 2002b, 1194)
- (4) a. They arranged [for the performance to begin at six].  
b. They intended [(for) the performance to begin at six].  
(Huddleston 2002b, 1178)
- (5) They expected [the performance to begin at six].  
[for excluded]  
(Huddleston 2002b, 1178)
- (6) a. We felt [the house shake]. (Lamprecht 1977, 253)  
b. I won't have [him criticize my work].  
(Lamprecht 1977, 254)  
c. They helped [me move the furniture].  
(Huddleston 2002b, 1174)
- (7) a. He says [(that) they are in Paris].  
(Huddleston 2002a, 951)  
b. We insist [that she be kept informed].  
(Huddleston 2002a, 993)

In this paper, I will focus on a specific aspect of some of these types of clausal embedding, namely referential dependencies between an argument of the main clause and the subject of the embedded complement clause. With infinitival clauses of the type in (3a) these dependencies have been discussed as control. My aim here is to consider central aspects of control by including referential dependencies across different types of propositional arguments in order to be able to distinguish the (lexical)-semantic aspects of control from the syntactic aspects of control, both of which restrict control. This implies that control results from the interaction of the lexical-semantic properties of the verb and the structural configuration. The main idea here is that in order to understand the **semantic input**, we need to look into referential dependencies in different realizations of propositional

arguments, not only infinitival complements. When investigating the **structural import** of control on the other hand, we need to consider cases in which the semantic restrictions of the verb are minimal. And if we do so across different languages, we will be able to differentiate the general mechanisms from the language specific properties of the syntactic realization of propositional arguments from other mechanisms. While this is a broader idea to follow, I will investigate the structural and (lexico-) semantic properties of control in two domains: (i) I report on the (lexico-) semantic properties of a small set of control verbs in German, and then, (ii) analyse the case of Dutch *zeggen*, where structural properties interact with meaning.

## 2. Background on control

In the literature on control, a major distinction has been made between those cases, in which the subject of an embedded infinitival clause needs to be co-referent with an argument of the matrix clause (obligatory control = OC) as in (8) and those examples in which this is not the case (non-obligatory control), see (9):

### (8) Obligatory control

- a. She<sub>i</sub> tried PRO<sub>i</sub> to be casual. (BNC, A08 2443)
- b. And he<sub>i</sub>'s promised PRO<sub>i</sub> to show us around. (BNC, A6B 1265)
- c. we persuaded him<sub>j</sub> PRO<sub>j</sub> to contact the [. . .] owners. (BNC, ACM 999)
- d. I asked him<sub>j</sub> PRO<sub>j</sub> to explain his pricing policy. (BNC, A14 730)

### (9) Non-obligatory control

- a. Clearly, [PRO confessing my crime] was not something they anticipated.
- b. I never understood why it is bad for health [PRO to stuff oneself with marshmallows].
- c. [After PRO pitching the tents], darkness fell quickly. (Landau 2013, 232)

While there is some disagreement as to the precise distinction, most researchers agree that these cases need to be kept apart in one way or another, as they have different properties, see Landau (2013) for a useful set of criteria to keep them apart. By and large, we find OC in complement clauses while NOC generally occurs in adjuncts and subjects. In the following, I will concentrate on the canonical cases of OC and put aside

both NOC as well as non-canonical cases in general, see Mucha et al. (in prep) for discussion of non-canonical cases.

There have been a range of proposals on how to account for the obligatory referential dependencies with respect to (i) the distinction between obligatory control and non-obligatory control, (ii) the regulations of controller choice and the possibilities of controller interpretation (e.g. exhaustive vs. partial control). Various researchers have made proposals that the underlying mechanism of control can be reduced to other mechanisms in grammar, such as – here I follow the grouping in Landau (2013) – Predication (see Williams 1980, Lebeaux 1984, Chierchia 1984), Binding (see Manzini 1983, Sag & Pollard 1991, Bouchard 1984, Koster 1984, Williams 1992, Manzini & Roussou 2000, for LFG see Bresnan 1982), A-movement (among others see Hornstein 1999, Polinsky & Potsdam 2002, Boeckx & Hornstein 2004, Manzini & Roussou 2000) or Agree (see among others Landau 2000 et seq.).

Besides the lively debate about the underlying syntactic mechanism of control, there is a range of approaches that suggest that the major ingredient in control and controller choice lies in the semantics of the verbs (see e.g., Köpcke & Panther 1993, 2002, Jackendoff & Culicover 2003, Culicover & Jackendoff 2005). This seems true on first sight for a number of control verbs which predominantly occur with infinitival complements, however, one major distinction that is rarely made and emphasized as relevant is that many verbs that restrict controller choice with infinitival clauses do not require the same referential dependency in clauses with overt subjects (finite, or overt subjects in infinitivals). This is true for example of the verbs *want* and *hope* in English, see (10) and (11).

- (10) a. John wanted PRO to leave.  
 b. John wanted for Mary to leave.
- (11) a. I hate to go and leave you in this state. (BNC, A0L 2382)  
 b. She hates that I have to fight against her countrymen. (BNC, CMP 1085)

Stiebels (2007, 2010) makes such a distinction of inherent vs. structural control, where **inherent control** requires co-reference of a specific argument of the matrix verb with the subject of the embedded constituent in all contexts, i.e., with subjects of non-finite (covert) and finite clauses (overt) and with subjects in nominalizations (overt or covert). The contrast is illustrated in (12) vs. (13) for German. **Structural control** on the other hand only requires co-reference with non-finite complements; with finite clauses, the reference of the embedded subject is free.

(12) Inherent Control: *ermutigen* ‘encourage’

- a. Maria<sub>i</sub> ermutigt ihren Sohn<sub>j</sub> [ <sub>]/\*/*ak*</sub> am Rennen  
 Maria encourages her.ACC son at-the race  
 teilzunehmen].  
 part-to-take  
 ‘Maria encourages her son to take part in the race.’
- b. ?Maria<sub>i</sub> ermutigt ihren Sohn<sub>j</sub>(dazu) [dass er<sub>j</sub>/*\*k*  
 Maria encourages her.ACC son thereof that he  
 am Rennen teilnimmt].  
 at.DEF race part-takes
- c. Maria<sub>i</sub> ermutigt ihren Sohn<sub>j</sub> [zur  <sub>]/\*/*ik*</sub> Teilnahme  
 Maria encourages her.ACC son to.DEF participation  
 am Rennen].  
 at.DEF race

(Stiebels 2010, 392, my gloss)

## (13) Structural Control

- a. Maria<sub>i</sub> hofft [ <sub>]/\*/*j*</sub> beim Rennen zu siegen].  
 Maria hopes at.DEF.DAT race to win
- b. Maria<sub>i</sub> hofft, [dass sie<sub>i</sub>/ <sub>]/\*/*j*</sub>Peter beim Rennen siegt].  
 Maria hopes that she/P. at.DEF.DAT race wins
- c. Maria<sub>i</sub> hofft auf ihren<sub>i</sub>/ <sub>]/\*/*j*</sub>Peters Sieg.  
 Maria hopes on her.ACC/P.s victory

(Stiebels 2010, 392, my gloss)

Taking this distinction seriously, we need to distinguish between those cases in which control is guided by the semantic properties of the selecting verb possibly interacting with the semantic properties of the complement, and those cases in which control is induced by the structure. In order to understand the syntactic input and structure of control, we therefore need to consider those verbs that do not give rise to inherent control. In turn, if we want to understand the semantic input, we need to provide a formal account of what exactly the semantic co-reference requirement is, and whether or not this is a lexical semantic property of the verb. In some cases, it might be difficult to keep the two apart.

With this much background, I want to look at two case-studies, one in which the semantic properties are decisive, anti-control in German, and another case where structural properties are more prominent, namely with the Dutch verb *zeggen*.

### 3. Lexical-semantic properties: Anti-Control in German

Wöllstein (2015), Brandt et al. (2016), Rapp et al. (2017), Brandt & Bildhauer (2019) analysed the selectional properties of a range of different control verbs and isolated a new class which they labeled anti-subject control verbs. These are verbs such as *missbilligen* ('to disapprove'), *würdigen* ('to appreciate') or *anordnen* ('to order/mandate'). The properties of these verbs are that (i) they occur with infinitival clauses only rarely, but when they do, (ii) they do not give rise to subject control even though the subject is the only available controller, and (iii) they only construe incoherently, which can be analysed as selection of a CP that blocks control. An example is given in (14).<sup>1</sup>

- (14) Maxi ordnet an, PRO\*<sub>ij</sub> die Zigaretten zurückzubringen.  
 Max order.3P.SG.PRS PRO DEF cigarette back.to.bring.INF  
 "Max orders (so.) to bring back the cigarettes."

In Hartmann & Mucha (2019), we take up this work and show that this anti-subject control property not only holds when these verbs combine with infinitival clauses: co-reference is also restricted with finite complements (and nominalizations):

- (15) Maxi ordnet an, dass er\*<sub>ij</sub> die Zigaretten  
 Max order.3P.SG.PRS that he DEF cigarettes  
 zurückbringt.  
 back.bring.3P.SG.PRS  
 "Max mandates that he bring back the cigarettes."

We take this to mean that anti-subject control is an inherent property of these verbs, which we formulate as a restriction of co-reference with the agent argument of the embedded verb (see Hartmann and Mucha 2019 for details). This semantic property is reflected in the structural properties: anti-subject control verbs are only compatible with CPs that are structurally large enough to block a syntactic mechanism of control, i.e. they either select finite clauses or construe necessarily incoherently.

## 4. Case study: Dutch *zeggen*

### 4.1 Introduction

Dutch *zeggen* ‘say’ falls in the class of verbs of communication (also termed as ‘illocutionary verbs,’ ‘speech act verbs’ or ‘verbs of saying,’ see Proost 2006). These are verbs that report a speech act and can be classified according to the original speech act, more specifically, the speakers’ attitude to the propositional content of the utterance reported. A directive verb of communication such as *order* describes a speech act in which the speaker wants the hearer to bring about the situation that is described in the original utterance. A commissive verb such as *allow* expresses that the speaker gives the hearer permission to bring about the situation that is described in the original utterance. A promissive verb such as *promise* describes that the speaker committed themselves to bringing about the situation that is described in the original utterance. A reportative verb provides the information that the speaker stated the proposition of the original utterance (possibly in a specific manner).

*Zeggen* can appear with both finite and infinitival clauses. When it occurs with infinitival clauses, *zeggen* is polysemous and can appear with a reportative reading and a directive reading illustrated in (17).<sup>2</sup>

- (16) a. Ze zei dat iedereen haar haatte  
 she said that everyone her hated  
 ‘She said that everyone hated her.’  
 (LASSY groot, wik\_part0601::1324135-12-7)
- b. Ze zei dat ze naar Salem moest komen  
 She said that she to Salem must.PAST come  
 ‘She said that she had to come to Salem.’  
 (LASSY groot, wik\_part0599::1318425-38-3)
- (17) a. Hij zei liberaal te zijn. REPORTATIVE  
 He said liberal to be  
 ‘He said that he is liberal.’  
 (LASSY groot, wik\_part0291::442480-10-3)
- b. Die zegt hem contact op te nemen  
 DEM.F.SG says him contact up to take  
 met de fotografe. DIRECTIVE  
 with the photographer  
 ‘She told him to contact the photographer.’  
 (LASSY groot, wik\_part0133::125018-15-7)

When the clause is finite, there is neither a requirement of co-reference nor a restriction of co-reference, see (16). Thus, this means that the verb *zeggen* does not encode control properties in its lexical entry. The investigation of the syntactic properties of this verb in combination with the infinitive is interesting as from a semantic-pragmatic perspective the directive reading seems to require control by the addressee on first sight, whereas no such restriction is expected with the reportative reading. In the following I show that the two readings correlate with a number of syntactic differences, which will lead to an analysis where the syntactic properties are decisive for the polysemy with this verb.

## 4.2 Syntactic aspects of *zeggen*

*Zeggen* in the reportative reading in Dutch can occur with both finite and non-finite complements, see (18).

- (18) a. In het artikel zegt hij echter geen communist  
 In DEF article says he really no communist  
 te zijn,  
 to be  
 (LASSY Groot, wik\_part0346::570808-9-2)
- b. In een interview zei hij dat hij nog steeds  
 in an interview said he that he still always  
 iedere dag gitaar speelde,  
 every day guitar played  
 (LASSY Groot, wik\_part0627::1400856-2-10))

When an infinitival clause is selected this is a semi-transparent *te*-infinitival clause; *om*-infinitivals are excluded (very clearly in the reportative reading, see below for details with respect to the directive reading). Semi-transparent infinitivals allow for arguments in the embedded clause to occur in the middlefield as in (19a), where the argument *een sterke affiniteit . . .* of the verb occurs in the middle field, while there is no verb cluster formation that gives rise to the IPP (infinitiv-pro-participium) effect, see (19b).<sup>3</sup>

- (19) a. die persoonlijk een sterke affiniteit met het  
 REL personally a strong affinity  
 with ..... his  
 land ..... zegt te voelen.  
 country ..... says to feel  
 ‘who says to feel a strong affinity with his country’  
 (LASSY Groot, wik\_part0105::85536-43-2)

- b. dat hij dit heeft {gezegd/\*zeggen} te voelen

Control with *zeggen* is subject to a strict licensing condition with the reportative reading, i.e., it requires an overt controller in the matrix clause. This can be seen with passive formation. In the reportative reading it is impossible to passivize *zeggen* as in (20a) (Broekhuis & Corver 2015). This licensing restriction is less strict in the directive reading, see (20b), which is possible with passive formation (see van Haaften 1991, 78) (the contrast is made clear here as the content of the embedded clause in (20a) is rather incompatible with a directive reading). With a directive reading, the addressee can remain implicit.

- (20) a. \*Er is Kees door Piet gezegd [PRO niet gelukkig te  
EXPL is K. by Piet said PRO not happy to  
zijn met die oplossing]  
be with the solution
- b. ?Er is ons (door moeder) gezegd [PRO vroeg thuis  
EXPL is us by mother said PRO early home  
te komen]  
to come

(van Haaften 1991: 78)

From a semantic-pragmatic perspective the restriction observed in (20a) vs. (20b) is difficult to handle: why should a demoted or absent argument be a viable antecedent for PRO in the directive reading but not in the reportative reading? Additionally, this cannot be a difference between subject (reportative) and object control (directive), as subject control verbs like *beloven* ‘promise’ also allow for the impersonal passive in Dutch, see (21).

- (21) Er werd ons<sub>j</sub> beloofd [(om) PRO<sub>arb</sub> de auto te  
there was us promised COMP PRO the car to  
repareren].  
repair

(Broekhuis & Corver 2015: 800)

I conclude from this restriction on reportative *zeggen* that the syntactic environment has an influence on the interpretation of PRO. As a working hypothesis, I would like to suggest that the two different readings of *zeggen* are not the result of two different lexical entries, but that the meaning of *zeggen* interacts with the structure (see below for more details). In order to establish the properties of the two readings, I probe into their control properties in the following sections.<sup>4</sup>



### 4.3 Predicative vs. logophoric control

#### 4.3.1 Overview

Landau (2015) distinguishes between two types of control, namely predicative vs. logophoric control, the two classes that were considered as partial control (=PC) vs. exhaustive control (=EC) (see Landau 2000, 2015, Pearson 2012, 2016 for more discussion); the two different types differ with respect to the properties given in table 1 taken from Landau (2015, 65).

	Predicative control	Logophoric control
Inflected Complement	yes	no
[-human] PRO	yes	no
Implicit control	no	yes
Control shift	no	yes
Partial control	no	yes
Split control	no	yes

**Table 1: Summary of empirical contrast between two types of control (Landau 2015: 65)**

As Dutch does not have inflected vs. non-inflected infinitives, I put this criterion aside and turn to the other criteria in turn.

#### 4.3.2 [-human] PRO

The criterion of [-human] cannot be tested easily independently of control, as the core meaning of *zeggen* in both readings prefers [+human] antecedents. This is certainly true for the crucial case of the directive reading disallowing [-human] antecedents for PRO, since the addressee for the directive reading, i.e., the antecedent has to be an entity that is able to bring about the proposition described in the embedded clause. For the reportative reading a [-human] antecedent seems to be possible as the following example shows:<sup>5</sup> The following example supports the idea that the reportative reading allows for a [-human] antecedent, in contrast to the directive reading.

- (22) De overeenkomst<sub>i</sub> zegt [ PRO<sub>i</sub> te eindigen op het moment dat de werkzaamheden naar het oordeel van opdrachtgever zijn voltooid.]  
 The contract says PRO to end at the moment that the work according to the judgment of customer are fulfilled  
 ‘The contract says that it ends as soon as the work is considered done by the customer’

Note that this kind of non-human antecedent falls in the class of sources, which can frequently replace human antecedents with a metaphorical shift. The crucial point though is that this metaphorical shift is not available for PRO in the embedded clause. No such cases seem possible with the directive reading.

### 4.3.3 Implicit control

As observed previously, the reportative reading does not allow for implicit control, see (23a), while the directive reading does, see van Haaften (1991), illustrated in (23b).

- (23) a. \*Er is Piet gezegd [ PRO Niet gelukkig te zijn met  
EXPL is Peter said PRO not happy to be with  
die oplossing]  
the solution

Literally: It was told to Peter not to be happy with the solution'

(van Haaften 1991, 78, my gloss and translation)

- b. En moeder heeft nog zo gezegd [ PRO op te passen  
and mother has PRT PRT said PRO up to pass  
voor mannen met een baard]  
for men with a beard  
'And Mom has said that we should watch out for men with a beard'

(van Haaften 1991, 79, my gloss and translation)

Additionally, a PP argument cannot be the controller in the reportative reading, while the directive reading is less marked:

- (24) a. \*Er is Kees door Piet gezegd [PRO niet  
EXPL is K. by Piet said pro not  
gelukkig te zijn met die oplossing]  
happy to be with the solution

- b. ?Er is ons door moeder gezegd [ PRO vroeg thuis  
EXPL is us by mother said PRO early home  
te komen]  
to come

(van Haaften 1991, 78)

This is confirmed by my native speaker informants: they do not accept the reportative reading with an implicit argument (the dropped *by*-phrase of

the passive) as in (25); the addressee argument is retained in the impersonal passive form.

- (25) De kinderen wordt gezegd [PRO het woonhuis te hebben  
 The children was said PRO the house to have  
 verkocht]  
 sold  
 ‘Literally: The kids were told to have sold the house.’

Thus, implicit control is possible with the directive reading, but not with the reportative reading.

#### 4.3.4 Control shift

The data for control shift on a descriptive level for the two readings needs to take into consideration the conceptual-semantic structure of the two readings, as well as the question whether or not control shift requires a “trigger” in the embedded infinitive. First, in the directive reading, there is a strong preference for control by the addressee, whereas nothing in the reportative situation requires such a restriction. As expected, the directive reading shows a strong preference for control by the addressee and control shift seems hardly possible see (26), even with a trigger in the embedded clause. Thus, *zeggen* in its directive reading seems to fall in the class of the so-called “verbs of influence” (term by Rooryck 2000) or implicative causative verbs (term by Landau 2015: 4), so control shift is not possible with the directive reading.

- (26) \*De studenten<sub>j</sub> zeiden de directeur<sub>i</sub> [ PRO<sub>j</sub> de activiteiten  
 DEF students said DEF director PRO DEF activities  
 te mogen continueren  
 to be.allowed continue  
 ‘Intended: The students told the director that they want to be allowed to continue their activities.’

This contrasts with the availability of control shift with *verzoeken* ‘ask’, a verb which should be rather similar in its semantic structure, see (27).

- (27) De studenten<sub>j</sub> verzochten de directeur<sub>i</sub> [PRO<sub>j</sub> de  
 DEF students continue asked DEF director PRO t DEF  
 activiteiten te mogen continueren.  
 activities to be.allowed  
 ‘The students asked the director to be allowed to continue their activities.’

For the reportative reading judgments for the availability of both subject and object vary. Most speakers only allow for the subject reading, see (28), but a few speakers also accept both readings.

- (28) Maria<sub>i</sub> zei haar vriendin<sub>j</sub> [PRO<sub>i</sub>%<sub>j</sub> zwanger te zijn].  
 Mary said her friend.FEM PRO pregnant to be  
 ‘Mary told her friend to be pregnant’

Note that it is striking that speakers require subject control, even though nothing in the reportative meaning forces such a reading.

#### 4.3.5 Partial control

The directive reading certainly allows for partial control readings in the following contexts. The reportative reading is judged more marked but probably marginally possible by my native speaker informants.

- (29) Directive: Jan<sub>i</sub> and Piet<sub>j</sub> are school children. After class they get into a fight. The teacher, Ms Harris, tries to settle the dispute and talks to the boys individually. She tells them how to interact in future.
- a. Zij zegt Piet<sub>j</sub> [PRO<sub>j+i</sub> in toekomst naar elkaar te luisteren].  
 She says P. PRO in future after each.other to listen  
 ‘She tells Piet that they should listen to each other in future.’
- b. Zij zegt Jan<sub>i</sub> [PRO<sub>i+j</sub> in toekomst met elkaar te praten].  
 She says Jan PRO in future with each.other to talk  
 ‘She tells Jan that they should talk to each other in future.’
- (30) Reportative Context: Jan<sub>i</sub> and Piet<sub>j</sub> are school children. After class they get into a fight. The teacher, Ms Harris, tries to settle the dispute and talks to the boys individually. She asks them individually about what they think the problem was.
- (31)
- a. ?Piet<sub>j</sub> zei [PRO<sub>j+i</sub> niet naar elkaar geluisterd te hebben].  
 P said PRO not after each.other listened to have  
 ‘Peter said that they did not listen to each other.’
- b. ?Jan<sub>i</sub> zei [PRO<sub>i+j</sub> niet met elkaar gepraat te hebben].  
 Jan said PRO not with each.other talked to have  
 ‘Jan said that they had not talked to each other.’

There is an additional issue as to what extent partial control is possible in transparent constructions with reportative *zeggen* as in (31). My informants report that the partial control reading forced by *elkaar* is rather impossible to get. Note though that it is not clear to what extent this is a cumulative effect of a marked word order and the partial control reading or whether partial control is not possible in general with the third construction, a topic which I leave to future research here, as it requires an in-depth experimental study independently of the verb *zeggen*.

- (32) a. \*Zij vertelde dat Piet naar elkaar zei  
 she told that P. after each.other said  
 geluisterd te hebben.  
 listen to have  
 ‘She told (me) that Piet said to have listened to each other.’  
 b. ??Zij vertelde dat Piet zei naar elkaar geluisterd te hebben.

### 4.3.6 Intermediate Summary

The preceding discussion has shown that the distribution of the different readings does not pattern along the lines of the two types of control as proposed in Landau (2015):

	Pred	Logoph.	Reportative	Directive
[-human] PRO	yes	no	yes	no
Implicit control	no	yes	no	yes
Control shift	no	yes	yes	hardly possible
Partial control	no	yes	marginally yes	yes

**Table 2: Comparison of reportative / directive reading with predicative / logophoric control**

However, on closer inspection, there is enough reason to suggest that the directive reading belongs to the class of logophoric control, whereas the reportative reading belongs to the class of predicative control.

So, let me consider the directive reading first, which patterns with logophoric control, with the exception of unavailability of control shift. This might actually not be so surprising, if, as will be discussed below in Section 4.4, the directive meaning is indeed not inherent in the meaning of the verb *zeggen*, but depends on the presence of a directive addressee and a future-oriented infinitive. The interpretation of PRO in this infinitive is in principle flexible, but the directive reading is closely linked to the addressee being

the agent of the embedded clause.

Turning to the reportative reading, the strongest argument for considering it a case of predicative control is the fact that it does not allow for implicit control. On the other hand, the marginal availability of partial control as an option in the reportative reading is not compatible with a strict view. As reported above, intuitions are rather tricky here and worth more in-depth discussion. To what extent speakers accommodate the intended reading in context that supports this reading exclusively can only be investigated with an experimental study along the lines of Pitteroff et al. (2017), which I leave to future research for the time being.<sup>6</sup>

The second possible argument against predicative control, is the availability of control shift. Note though that control shift is dependent on the overt presence of the respective argument, so that possibly both structures are available for the verb *zeggen*, see below.

#### 4.4 Control and clausal embedding: a tentative proposal

Summarizing the discussion so far, we have seen the following:

- the directive reading correlates with the appearance of an opaque *om-te*-infinitive;
- the reportative reading correlates with a transparent *te*-infinitive;
- the directive reading shows most of the hallmark of logophoric control;
- the reportative reading shows the main hallmarks of predicative control;

Both readings use the same verb, so we need to figure out whether or not we should assume two different meanings of the verb with a different argument structure or whether there is a common core meaning of the verb which can combine with different arguments to give rise to the different readings.

In the following I sketch a proposal that allows for a shared meaning of *zeggen* to combine with different constituents. The core idea is that the directive reading needs to select an infinitival complement that marks a directive speech act giving rise to a fully projected CP, while the reportative reading selects for a smaller constituent, a TP, presumably because it can rely on the reportative reading to be a default that is not required to be marked in the infinitival complement.

#### 4.4.1 Basic assumptions

In this section, I spell out the background for my analysis. First, in infinitival complement clauses in Dutch the subject argument is unpronounced. I take this element to be UPRO (see McFadden & Sundaresan 2016, 2018), i.e., a pronominal form whose referential properties depend on the syntactic environment it occurs in. In combination with the analysis of control in Landau (2015) we can get the following patterns of control in Dutch: UPRO can either be directly bound/agreed with by an argument in the matrix clause under *c*-command, when the complement clause does not introduce a C-layer that blocks direct agreement (this corresponds to Landau's 2015 predicative control). For the cases discussed here, I assume that this C-layer is just absent with transparent infinitives, though it is conceivable that the C-layer is just not of the type that blocks direct agree relations.

Second, I assume that UPRO can also be bound/agreed with indirectly via the C-layer (some kind of logophoric center) which in turn gets its referential properties from an argument in the matrix clause (this is the logophoric control in Landau's approach). In Dutch, the C-layer can be made overt by the use of the complementizer *om*.<sup>7</sup>

Besides these basic assumptions about Control, I follow Grimshaw (2015, 2017) in that the communication verbs include a light verb SAY which adds the two syntactic frames given in (32) where the second is compatible with [-human] subjects.

- (33) a. SAY1 {Agent/*i* Linguistic-Material/*j* Goal/*k*}  
 b. SAY2 {Location/*i* Linguistic-Material/*j*}

The two schemas in (32a) and (32b) additionally differ with respect to the availability of a goal argument, which is possible (though not obligatory) in (32a), but absent with SAY2.

Additionally, the schema for SAY1 is assumed to be quite broad in Grimshaw (2017) in the sense that *say*-verbs report speech events in general, but the type of speech event is encoded in the 'linguistic material'-argument; some verbs restrict the type of speech act—*ask* only selects questioning speech acts—others are free of such restrictions—for example *mutter* can occur with both *wh*- and *that*-clauses.

Grimshaw (2015, 2017) does not discuss directive speech acts. Nevertheless, her proposal can be expanded quite naturally to include these cases by extending the range of force types that the ‘linguistic material’-argument can encode including D-Force (directive force), the third option besides A-force (assertive force) and Q-force (questioning force).

- (34) a. He said to me ‘Leave!’.  
 b. ‘Leave!’, he said.

As the directive reading requires an agent and a goal (addressee) argument, this is only possible with the first frame of the light verb in (32a).

Additionally, we need to add that D-force complements can in principle be infinitival and indeed seem to be generally so in Dutch; as force is usually considered a C-related projection we expect D-force infinitivals to have C-properties. In Dutch, this is the overt C-element *om*.

Turning now to the second scheme, Grimshaw (2015) assumes that the hall-mark for the SAY2 scheme is that the matrix argument is inanimate providing the location/source of the information reported in the linguistic complement, see (34). Additionally, no addressee argument is possible with this frame.

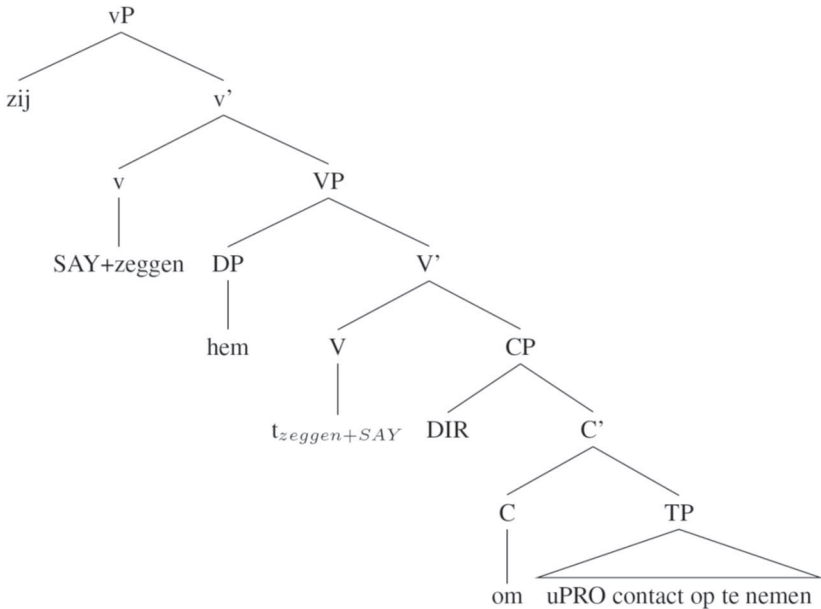
- (35) The poster said that the park was closed. (Grimshaw 2015, 86)

#### 4.4.2 The directive reading

With this much as background, we can now propose an analysis for directive *zeggen*. For the directive reading, the crucial ingredient in the analysis proposed here is that it arises if the complement is a full-fledged CP, which specifies the directive speech act as just discussed above. As the directive reading relies on a full-fledged CP, and as the nature of direction is addressee-oriented, it gives rise to logophoric control with the addressee as the obligatory antecedent. The full structure is given in (35).<sup>8</sup>



- (36) Zij zegt hem UPRO contact op te nemen



#### 4.4.3 The reportative reading

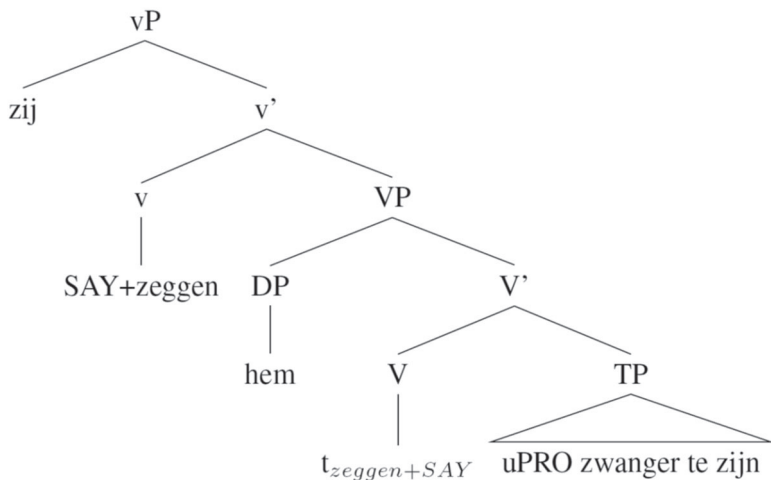
Within the reportative reading, I think that we need to distinguish at least the two different patterns introduced by Grimshaw, see (32a) and (32b). Most reportative cases seem to be cases of SAY1 with an optional addressee argument. Additionally, we find the SAY2 pattern with [-human] subjects and lacking an addressee argument, see the examples in (23) above. In these cases, the subject is the location of the information.<sup>9</sup>

We often find SAY1 with finite clauses for *zeggen* in Dutch, see (16) above, but we also find *zeggen*+INF with an additional argument, even though these are not frequent with *zeggen*, see (36).<sup>10</sup>

- (37) De vader Hildebrand      zegt    zijn zoon      nooit een verwant  
 DEF father Hildebrand    says    his son        never a relative  
 te hebben gehad, [...]  
 to have    had  
 ‘Father Hildebrand told his son that he never has had a relative.’  
 (Lassy groot: wik\_part0069::42257-17-4)

However, these declarative infinitivals do not project a full CP, see (37), but are interpreted as declaratives as default.<sup>11</sup>

(38) Zij zegt hem UPRO zwanger te zijn



The basis for assuming a TP with the reportative reading is the fact that implicit control is not possible. Consider first, the ungrammaticality of the passive form in (38):<sup>12</sup>

(39) \*Er wordt gezegd [ PRO niet gelukkig te zijn met  
 EXPL AUX said [ PRO not happy to be with  
 die oplossing]  
 the solution]

The ungrammaticality is a result of the TP structure: a passive form of the reportative reading of the verb *zeggen* results in a raising configuration, PRO is required to move to the subject position where UPRO is not licensed as silent pronominal; it can only appear as an overt pronoun, as in (39).

(40) Hij wordt gezegd [ PRO niet gelukkig te zijn met die oplossing]  
 He AUX said [ not happy to be with the solution]

As *zeggen* does not select a CP complement, UPRO cannot receive a logophoric or pronominal interpretation via co-reference with an implicit argument. As a result, we get a configuration of obligatory structural control

which is not required/forced by the verb, but merely by the structural configuration. This is a case that any analysis that reduces control exclusively to verb semantics cannot handle.

## 4.5 Summary

Summarizing the discussion on Dutch *zeggen*, I have proposed that directive *zeggen* selects for a directive CP complement, i.e. the directive meaning is a result of the combinatorics in the structure, whereas reportative *zeggen* occurs with a TP only. Furthermore, I suggested that reportative *zeggen* combines with SAY1 or SAY2, in the latter case giving rise to non-human antecedents.

## 5. Conclusion

In this paper I have argued that both lexical-semantic and structural properties interact in giving rise to control or anti-control. I have briefly presented evidence for a lexical-semantic analysis of control concerning anti-subject-control verbs in German. For the polysemous Dutch verb *zeggen*, I have argued that the polysemy is a result of a combination of a general light verb SAY1 with either a directive CP (which can host *om*) or a reduced TP which gives rise to obligatory control even though the verb meaning does not require such a co-reference restriction.

## References

- Beek, Janneke ter. 2008. *Restructuring and infinitival complements in Dutch*, vol. 68 LOT dissertation series. Utrecht: LOT.
- Boeckx, Cedric & Norbert Hornstein. 2004. "Movement Under Control". *Linguistic Inquiry* 35(3): 431–452.
- Bouchard, Denis. 1984. *On the Content of Empty Categories*. Dordrecht: Foris.
- Brandt, Patrick & Felix Bildhauer. 2019. "Alternation von zu- und dass-Komplementen: Kontrolle, Korpus und Grammatik". In *Grammatik im Korpus*, edited by Marek Konopka and Angelika Wöllstein, 211–298. Tübingen: Narr.
- Brandt, Patrick, Beata Trawin'ski & Angelika Wöllstein. 2016. "(Anti-)Control in German: Evidence from comparative, corpus- and psycholinguistic studies." In *Co-and Subordination in German and other Languages, Linguistische Berichte, Sonderheft 21*, edited by Ingo Reich, and Augustin Speyer, 77–98. Hamburg: Buske.

- Bresnan, Joan. 1982. "Control and complementation." *Linguistic Inquiry* 13 (3): 343–434.
- Broekhuis, Hans, Hans den Besten, Kees Hoekstra and Jean Rutten. 1995. "Infinitival complementation: On remnant extraposition." *The Linguistic Review* 12 (2): 93–122.
- Broekhuis, Hans & Norbert Corver. 2015. *Syntax of Dutch: Verbs and Verb Phrases: Volume 2* Comprehensive Grammar Resources. Amsterdam: Amsterdam University Press.
- Chierchia, Gennaro. 1984. "Topics in the Syntax and Semantics of Infinitives and Gerunds." PhD diss. University of Massachusetts.
- Cinque, Guglielmo. 2006. *Restructuring and Functional Heads. The Cartography of Syntactic Structures*. Oxford Studies in Comparative Syntax. Oxford: Oxford University Press.
- Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler Syntax*. Oxford and New York: Oxford University Press.
- den Besten, Hans, Jean Rutten, Toen Veenstra & Joop Veld. 1988. "Verb raising, extraposition en de derde constructie [Verb raising, extraposition and the third construction]" Ms., University of Amsterdam.
- Grimshaw, Jane. 2015. "The light verbs say and SAY." In *Structures in the mind*, edited by Ida Toivonen, Piroska Csúri and Emile van der Zee, 79–100. Cambridge, Massachusetts and London, England: The MIT Press.
- Grimshaw, Jane. 2017. "Force compatibility explains 'selection' patterns exhibited by SAY-verbs.", paper presented at SelectionFest ZAS 10/11/2017.
- Hartmann, Jutta M. 2018. "Control and the Dutch polysemous verb 'zeggen'." Ms., IDS Mannheim.
- Hartmann, Jutta M. and Anne Mucha. 2019. "Restricting Control: The Role of Syntax and Semantics in Anti-Control.", Ms., IDS Mannheim.
- Hoekstra, Teun. 1983. "The distribution of sentential complements." In *Linguistics in the Netherlands*, edited by Bennis, Hans and W.U.S. van Lesser Kloeke, 93–103. Dordrecht: Foris.
- Hornstein, Norbert. 1999. "Movement and Control." *Linguistic Inquiry* 30 (1): 69–96.
- Huddleston, Rodney D. 2002a. "Content Clauses and Reported Speech." In *The Cambridge Grammar of the English Language*, edited by Rodney D. Huddleston, Geoffrey K. Pullum & Laurie Bauer, 947–1030. Cambridge: Cambridge University Press.
- Huddleston, Rodney D. 2002b. "Non-finite and Verbless Clauses". In *The Cambridge Grammar of the English Language*, edited by Rodney D. Huddleston, Geoffrey K. Pullum and Laurie Bauer, 1171–1271. Cambridge: Cambridge University Press.

- Jackendoff, Ray and Peter W. Culicover. 2003. "The Semantic Basis of Control in English." *Language* 79 (3): 517–556.
- Köpcke, Klaus-Michael and Klaus-Uwe Panther. 1993. "A Cognitive Approach to Obligatory Control Phenomena in English and German." *Folia Linguistica* 17 (1-2): 57–105.
- Köpcke, Klaus-Michael and Klaus-Uwe Panther. 2002. "Zur Identifikation leerer Subjekte in infinitivischen Komplementsätzen – ein semantisch-pragmatisches Modell". *Folia Linguistica* 36 (3-4): 191-218.
- Koster, Jan. 1984. "On Binding and Control." *Linguistic Inquiry* 15 (3): 417–459.
- Lamprecht, Adolf. 1977. *Grammatik der englischen Sprache*. Berlin: Cornelsen.
- Mucha, Anne, Jutta M. Hartmann and Beata Trawiński (eds.). in prep. *Non-canonical control in a cross-linguistic perspective*. Amsterdam: Benjamins.
- Landau, Idan. 2000. *Elements of Control: Structure and Meaning in Infinitival Constructions*, vol. 51 Studies in Natural Language and Linguistic Theory. Dordrecht: Springer.
- Landau, Idan. 2013. *Control in generative grammar: A research companion*. Cambridge: Cambridge University Press.
- Landau, Idan. 2015. *A Two-Tiered Theory of Control* vol. 71, Linguistic Inquiry Monographs. Cambridge: The MIT Press.
- Landau, Idan. 2020. "Nonobligatory Control with Communication Verbs: New Evidence and Implications." *Linguistic Inquiry* 51 (1): 75–96.
- Lebeaux, David. 1984. "Anaphoric Binding and the Definition of PRO." *Proceedings of NELS* 14: 253–275.
- Manzini, Maria Rita. 1983. "On control and control theory." *Linguistic Inquiry* 14: 421–446.
- Manzini, Maria Rita & Anna Roussou. 2000. "A minimalist theory of A-movement and control." *Lingua* 110 (6): 409–447.
- McFadden, Thomas and Sandhya Sundaresan. 2016. "Failure to control is not failure: it's pro." In *NELS 46 Proceedings of the North East Linguistic Society (NELS)*, edited by Christopher Hammerly, and Brandon Prickett, 1-10. Amherst, MA: University of Massachusetts.
- McFadden, Thomas and Sandhya Sundaresan. 2018. "Reducing pro and PRO to a single source." *The Linguistic Review* 35(3): 463–518.
- Pearson, Hazel A. 2012. *The Sense of Self: Topics in the Semantics of De Se Expressions*. PhD diss., Harvard University.
- Pearson, Hazel A. 2016. "The semantics of partial control." *Natural Language & Linguistic Theory* 34 (2): 691–738.
- Pitteroff, Marcel, Artemis Alexiadou, Jeannique Darby, and Silke Fischer.

2017. "On partial control in German." *The Journal of Comparative Germanic Linguistics* 20 (2): 139–185.
- Polinsky, Maria & Eric Potsdam. 2002. "Backward Control." *Linguistic Inquiry* 33 (2): 245–282.
- Proost, Kristel. 2006. "Speech Act Verbs." In *Encyclopedia of language & linguistics*, edited by Keith Brown, 651–656. Amsterdam: Elsevier.
- Rapp, Irene, Ekaterina Laptieva, Alexander Kopenlig, and Stefan Engelberg. 2017. "Lexikalisch-semantische Passung und argumentstrukturelle Trägheit – eine korpusbasierte Analyse zur Alternation zwischen dass- Sätzen und zu-Infinitiven in Objektfunktion." *Deutsche Sprache* 45 (3): 193–221.
- Rooryck, Johan. 2000. *Configurations of sentential complementation: Perspectives from Romance languages*, vol. 4 Routledge leading linguists. London and New York: Routledge.
- Sag, Ivan A. and Carl Pollard. 1991. "An Integrated Theory of Complement Control." *Language* 67 (1): 63–113.
- Stiebels, Barbara. 2007. "Towards a typology of complement control." In *Studies in Complement Control ZAS Papers in Linguistics*, edited by Barbara Stiebels, 1–80. Berlin: ZAS.
- Stiebels, Barbara. 2010. "Inhärente Kontrollprädikate im Deutschen". *Linguistische Berichte. Hamburg* (224): 391–440.
- van Haaften, Ton. 1991. *De interpretatie van verzweggen subjecten*. PhD diss., Free University Amsterdam.
- Williams, Edwin. 1980. "Predication". *Linguistic Inquiry* 11 (1): 203–238.
- Williams, Edwin. 1992. "Adjunct Control". In *Control and Grammar*, edited by Richard Larson, Sabine Iatridou, Utpal Lahiri, and James Higginbotham, 297–322. Dordrecht: Kluwer.
- Wöllstein, Angelika. 2015. "Grammatik-explorativ". In *Sprachwissenschaft im Fokus* Jahrbuch des Instituts für Deutsche Sprache, edited by Ludwig Eichinger, 93–120. Berlin and New York: Mouton de Gruyter.
- Wöllstein-Leisten, Angelika. 2001. *Zur Syntax der dritten Konstruktion*. Tübingen: Stauffenberg.
- Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure*, vol. 55 Studies in generative grammar. Berlin and New York: Mouton de Gruyter.
- Zwart, Jan-Wouter. 1997. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*, vol. 39 Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer Acad. Publ.
- Zwart, Jan-Wouter. 2011. *The Syntax of Dutch*. Cambridge: Cambridge University Press.

## Notes

<sup>1</sup> Note that these cannot be cases of implicit control as *anordnen* cannot select for an addressee argument.

(i) \*Max<sub>i</sub> ordnet Paul an, die Zigaretten zurückzubringen.  
 Max order.3SG.PRS Paul PRT the cigarettes back.to.bring.INF  
 Intended: “Max orders Paul to bring back the cigarettes.”

<sup>2</sup> There might be an additional promissive reading which is not easy to distinguish from the reportative in some cases

<sup>3</sup> Broekhuis et al. 1995 analyses these semi-transparent *te*-infinitives as extraposition plus scrambling versus verb cluster formation; whether or not this is the correct analysis is not central to my concern here. The important point, as will be made clear below is that transparent *te*-infinitives are not fully clausal, i.e. they do not contain a C-layer, despite being extraposed. Transparent infinitives have also been discussed under the label Third Construction, German *Dritte Konstruktion*, see den Besten et al. (1988), Beek (2008) for discussion on Dutch and Wöllstein-Leisten (2001), Wurmbrand (2001) for discussion on German.

<sup>4</sup> Due to lack of space, I do not discuss whether or not non-obligatory control might be a relevant category in line with the discussion in Landau (2020). According to the criteria established in Landau (2013) (arbitrary control, long-distance control, discourse control, non-c-commanding control), both readings fall in the category of obligatory control. Note that most of my informants report that the translation of *Dad said to be quiet* is impossible in Dutch in a configuration where the person that Dad addresses is not the referent of the embedded clause, in contrast to what has been reported for English by Landau (2020).

<sup>5</sup> A [-human] antecedent might be established in the context of *wh* infinitives (see Landau 2015, 67), but *wh*-infinitives are marginal in Dutch to begin with (Zwart 2011)

<sup>6</sup> There is an additional related issue of whether or not the reportative reading with *zeggen* can be considered a propositional attitude verb. It certainly need not be as the availability of a [-human] antecedent as discussed in 4.3.2 shows.

<sup>7</sup> Note that UPRO can have pronominal properties resulting in non-obligatory control. I put this aside, see McFadden & Sundaesan (2018) for details.

<sup>8</sup> I abstract away here from the derivation of Dutch OV and extraposed infinitival clauses. Following Zwart (1997), I assume that extraposed clauses are base-generated to the right. They might also be base-generated to the left and extraposed to the right, see Hoekstra (1983).

<sup>9</sup> It is not obvious to me, why a [+human] subject might not be reported as the source of the information, resulting in an ambiguity of SAY1 and SAY2 when there is no addressee argument. Such an analysis would be compatible with SAY2 being some evidential projection in the clausal spine along the lines of Cinque (2006).

<sup>10</sup> In a corpus study on communicative verbs in Dutch, (see Hartmann 2018) I only found one such case out of 160 relevant cases with the reportative reading.

<sup>11</sup> This complement can in principle be also a C-related projection that is small

---

enough not to introduce its own logophoric domain. I put this option aside for simplicity sake.

<sup>12</sup> Note that such passives are possible with finite clauses.



# CHAPTER TWO

## SCOPE DEPENDENCIES WITH ROMANIAN DITRANSITIVES<sup>1</sup>

ALINA TIGĂU

Scope dependencies and the existence of scope freezing effects with ditransitives have been employed as a diagnostic test for the identification of the underlying order of the two internal arguments with these constructions. This paper reports on some experimental results regarding scope dependencies within Romanian ditransitives and shows that none of the possible configurations is restricted to one scope interpretation only, so that one cannot claim that they exhibit scope freezing effects as it is the case for other languages such as Russian, Japanese, or English. Nevertheless, some preferences regarding scope interpretation are uncovered, which function as evidence that the underlying configuration is the *Goal-above-Theme* one with Romanian ditransitives. Direct object marking also seems to be an important factor influencing scope dependencies with ditransitives.

Keywords: Ditransitives, scope, internal arguments, word order

### 1. Preliminaries

Scope dependencies within ditransitive constructions and the existence of scope freezing effects have long constituted a diagnostic test for the identification of the underlying order of the two internal arguments. Aoun and Li (1989) are among the first to draw attention to the fixed scope dependency holding between the two internal arguments within the Double Object Construction (DOC) in English. In (1) the indirect object (IO) *someone* outscopes the direct object (DO) *every problem* and this seems to be the only possible reading (an interpretation with a wide scope reading on *every problem* is out).

- (1) I assigned someone every problem.

Aoun and Li (1989): (11) p. 147

The same goes for the examples under (2), where the IO may only have a wide scope interpretation with respect to the DO:

- (2) a. The committee gave some student every book in the library.  
b. John asked two students every question.

Aoun and Li (1989): fn 23 (i) p. 166

This state of affairs has been argued to arise as a consequence of an asymmetric c-command relation which holds between the IO and the DO within the DOC, and whereby the IO is hierarchically superior to the DO. Binding dependencies reinforce this hypothesis: the IO seems to be able to bind into the DO while the opposite binding relation may not obtain. Barss and Lasnik (1986) discuss a number of phenomena in this respect:

*Anaphor binding*

- (3) a. I showed John<sub>i</sub> himself<sub>i</sub> in the mirror.  
b. \*I showed himself<sub>i</sub> John<sub>i</sub> in the mirror.

Barss and Lasnik (1986): (2), (3) p. 347

*Quantifier binding*

- (4) a. I denied each worker<sub>i</sub> his<sub>i</sub> paycheck.  
b. I showed every friend<sub>i</sub> of mine his<sub>i</sub> photograph  
c. \*I denied its owner each paycheck.  
d. \*I showed its trainer every lion.

Barss and Lasnik (1986): (6), (7) p. 348

*Wh-movement and Weak Crossover*

- (5) a. Which worker<sub>i</sub> did you deny his<sub>i</sub> paycheck?  
b. Who<sub>i</sub> did you show his<sub>i</sub> reflection in the mirror?  
c. \*Which paycheck<sub>i</sub> did you deny its<sub>i</sub> owner?  
d. \*Which lion<sub>i</sub> did you show its<sub>i</sub> trainer?

Barss and Lasnik (1986): (8), (9) p. 348

Unlike the DOC, where the two quantified internal arguments are scopally unambiguous, being limited to the surface scope, the Prepositional Object Construction (POC) allows for scope ambiguity. In (6a) below both the DO and the IO may acquire a wide scope interpretation as pointed out by Larson (1990). The DOC variant in (6b) is expectedly restricted to a wide scope reading on the IO:

- (6) a. The teacher assigned one problem to every student.  
 b. The teacher assigned one student every problem.

Larson (1990): (20), p. 604

Marantz (1993) argues that this difference between the DOC and the POC variants finds justification in their different underlying structure: while the DOC gives rise to a complex structure featuring an Applicative Projection (which hosts the IO in its specifier) above the VP (which contains the DO), in the POC the two internal arguments are to be found inside the same VP projection, which allows them both to undergo QR and thereby acquire a wide scope interpretation. In the DOC, on the other hand, the DO within the VP cannot undergo QR above the IO within the ApplP.

Other languages seem to display similar scope freezing effects with respect to their ditransitive constructions.

**Russian**, for instance, evinces various scope freezing effects with several classes of ditransitives<sup>2</sup>. In (7a), where the order of the two arguments is  $DP_{ACC} > DP_{DAT}$ , the sentence is ambiguous with respect to the scope of the two arguments. When the opposite surface word order obtains (7b), the only possible reading is one in which the dative argument has wide scope:

- (7) a. Maša        prostila kakoje-to        predatel'stvo  
 Masha        forgave some                    betrayal.ACC  
 každoi podruge.  
 every girlfriend.DAT  
 'Masha forgave some betrayal to every girlfriend.'
- b. Maša        prostila kakoi-to podruge  
 Masha        forgave some    girlfriend.DAT  
 každoje        predatel'stvo.  
 every            betrayal.ACC  
 'Masha forgave some girlfriend every betrayal.'

Antonyuk (2015): (12) p. 106

The same obtains in (8): when the IO precedes the DO (8b), the only interpretation available is one according to which the DO has wide scope, while the order DO > IO is ambiguous for scope (8a):

- (8) a. Učitel' predložil kakuju-to knigu  
 Teacher.NOM offered some book.ACC  
 kazdomu studentu  
 every student.DAT  
 'The teacher offered some book to every student.'
- b. Učitel' predložil kakomu-to studentu  
 Teacher.NOM offered some student.DAT  
 každuju knigu.ACC  
 every book.  
 'The teacher presented some student with every book.'

Antonyuk (2015): (57), (58) p. 83

In (9) it is the opposite word order which gives rise to scope freezing effects: the accusative object in (9b) may only be interpreted with wide scope, while the DP<sub>INSTR</sub> > DP<sub>ACC</sub> order is scopally ambiguous:

- (9) a. Maša ugostila kakim-to pečenjem každogo rebenka.  
 Masha treated some cookie.INSTR every child.ACC  
 'Masha treated every child to some cookie.'
- b. Maša ugostila kakogo-to rebenka každym pečenjem  
 Masha treated some child.ACC every cookie.INSTR  
 'Masha treated some child to every cookie.'

Antonyuk (2015): (13) p. 106

**Japanese** has also been argued to display scope freezing effects within ditransitive configurations. Hoji (1985) argues that in the DP<sub>DAT</sub> > DP<sub>ACC</sub> the IO asymmetrically outscopes the DO, while the opposite surface word order is ambiguous with respect to scope. Based on these facts, Hoji (1985) posits that the DP<sub>DAT</sub> > DP<sub>ACC</sub> is the basic configuration, while the DP<sub>ACC</sub> > DP<sub>DAT</sub> order is analyzed as derived by way of scrambling the DO over the IO. The DO may then reconstruct in its merge position, which accounts for its narrow scope interpretation being still available.

- (10) a. Taroo-ga dareka-ni dono-nimotu-mo okutta.  
 Taro.NOM someone.DAT every-package sent  
 ‘Taro sent someone every package.’  
 some > every, \*every > some
- b. Taroo-ga dono-nimotu-mi<sub>i</sub> dareka-ni t<sub>i</sub>  
 Taro.NOM every-package<sub>i</sub> someone.DAT t<sub>i</sub>  
 okutta.  
 sent  
 some > every, every > some

Miyagawa and Tsujioka (2004): 11, 12, p. 5

Miyagawa and Tsujioka (2004) refine the data in (10) by contributing the example in (11) below, which is characterized by scope ambiguity:

- (11) Taroo-ga dokoka-ni dono-nimotu-mo okutta.  
 Taro.NOM some place-to every-package sent  
 ‘Taro sent every package to some place.’  
 Some > every, every > some

Miyagawa and Tsujioka (2004): 13 p. 6

They distinguish between structures with an animate Goal DP, which they assimilate to the English DOC, and structures where the Goal is inanimate and where the same scope ambiguity as in the English POC obtains. The former structures are argued to evince a higher Goal position which will always c-command the position occupied by the Theme DP (and where the Goal is always interpreted as possessor). The structures paralleling (11), on the other hand, are argued to contain a lower Goal position (interpreted as Location), which may be either followed or preceded by the Theme DP – hence the scope ambiguity which obtains:

- (12) high Goal > (Theme) > low Goal > (Theme)

Scope freezing is thus argued to arise only with the DOC counterparts, where the IO occupies the higher Goal position.

In this paper we would like to report on an experiment we carried out on Romanian ditransitives with the purpose of uncovering whether any of the word orders at stake gives rise to scope freezing effects. Evidence along these lines, would allow us further insights into the underlying structure of these constructions.

Romanian allows for a *Goal-over-Theme* as well as a *Theme-over-Goal* configuration. Furthermore, both the IO and the DO may be clitic doubled:

- (13) a. Maria (i) (l-)a prezentat  
 Mary (him.3P.SG.DAT) (him.3P.SG.M.ACC)-has introduced  
 pe Ion lui Matei.  
 DOM John DAT Matthew  
 ‘Mary introduced John to Matthew.’
- b. Maria (i) (l-)a prezentat  
 Mary (him.3P.SG.DAT) (him.3P.SG.ACC)-has introduced  
 lui Matei pe Ion.  
 DAT Matthew DOM John  
 ‘Mary introduced John to Matthew.’

Tigău (2020) provides arguments in favour of a basic Goal-over-Theme configuration and, building on extensive experimental evidence, shows that the two internal arguments have symmetric binding potential. The current paper contributes further evidence in favour of this hypothesis by presenting the results of three experiments which probe for scope dependencies within ditransitives. As will be seen, unlike languages such as English, Russian or Japanese presented above, where certain ditransitive configurations are characterized by scope freezing effects, Romanian ditransitives seem to be ambiguous with respect to scope in both surface word orders and irrespective of whether either or both of the two internal arguments has/have been clitic doubled or not. This goes on a par with the symmetric binding potential of the two arguments uncovered in a series of independent experiments reported in Tigău (2020) and Cornilescu et al. (2017a,b).

This paper has the following structure: we begin by describing the basic tenets of the three experiments in section 2 and continue with a discussion of the results in 3. Section 4 contains the conclusions.

## 2. Experimental data on scope dependencies within Romanian ditransitives

This section dwells on three experiments carried out on ditransitive configurations with the aim to verify the scope relations holding between the two internal arguments. The experiments were similar in design and different only with respect to the type of DO: given that Romanian exhibits unmarked DOs, differentially marked DOs (DOMed DOs) and clitic doubled + differentially marked DOs (CDed+DOMed DOs), we felt it was necessary to distinguish between these three types of DO and to investigate scope dependencies between one DO type and the IO at a time. Consequently, **Experiment 1** focused on the scope interactions between unmarked DOs and IOs; **Experiment 2** dealt with scope dependencies obtaining with DOMed DOs and IOs, while **Experiment 3** dwelt on the scope relations obtaining between the two internal arguments when CDed+DOMed DOs were involved.

Apart from the DO type which was kept constant in each experiment, there were several parameters which we varied: the surface order of the two internal arguments with respect to each other, the doubling of the IO by means of a dative pronominal clitic or the lack thereof, the type of DP – indefinite vs. universal quantifier. Table 1 shows how these parameters interacted within each experiment:

DO before IO				IO before DO			
DO <sub>UQ</sub> bef. IO <sub>ind</sub>		DO <sub>ind</sub> bef. IO <sub>UQ</sub>		IO <sub>UQ</sub> bef. DO <sub>ind</sub>		IO <sub>ind</sub> bef. DO <sub>UQ</sub>	
-cl <sub>IO</sub>	+cl <sub>IO</sub>	-cl <sub>IO</sub>	+cl <sub>IO</sub>	-cl <sub>IO</sub>	+cl <sub>IO</sub>	-cl <sub>IO</sub>	+cl <sub>IO</sub>
8 items	8 items	8 items	8 items	8 items	8 items	8 items	8 items

**Table 1** experiment design

We started from 8 basic sentences built around different ditransitive verbs and then varied word order, clitic doubling and DP type as shown in Table 1, such that a total of 64 experimental items was obtained. The 64 items were then assigned to four different lists using the Latin square method. To each list we then added 14 fillers and the 30 items in each list were afterwards randomly ordered and entered into Google forms. The resulting 4 questionnaires were then randomly assigned to native speakers of Romanian, mostly students of the University of Bucharest with ages ranging from 20 to 24. Each questionnaire was assessed by at least 20 native speakers of Romanian so that more than 80 respondents took part in each

experiment, with an overall participant score of over 240 respondents for the entire experimental undertaking.

The respondents had to perform two tasks for each experimental item:

- a) **An acceptability task:** the participants had to assess the respective item with respect to its acceptability by assigning it values ranging from 1 to 7 on an acceptability scale where 1 represented the lowest acceptability score and 7, the highest.
- b) **A norming task:** the respondents had to select as possible readings among the following interpretations:
  - i. the wide scope interpretation for the indefinite DP
  - ii. the narrow scope interpretation for the indefinite DP
  - iii. both readings

Here is an actually tested item together with the accompanying norming task:

- (14) Încolțit de armatele inamice, împăratul a trimis  
 Cornered by armies.the enemy, emperor.the has sent  
 un mesager fiecărui aliat pentru a cere ajutor.  
 a messenger.ACC every.DAT ally for to ask help  
 ‘Cornered by enemy armies, the emperor has sent a messenger to every ally in order to ask for help.’

Task: Please pick the possible interpretation(s) for the sentence above:

- a) The emperor sent to each of his allies a different messenger
- b) The emperor sent the same messenger to each of his allies
- c) Both (a) and (b)

As already mentioned, each experimental item was modified by changing the relative order of the two internal arguments with respect to each other, their DP type (i.e., indefinite vs. universal QP), and the presence/absence of a dative pronominal clitic doubling the IO. Each variant was accompanied by its corresponding interpretation possibilities. Example (15) contains all the variants of (14) above:



(15) **DO<sub>UQ</sub> before IO<sub>ind</sub>, no IO clitic**

- a. Încolțit de armatele inamice, împăratul a  
 Cornered by armies.DEF.PL enemy.PL, emperor.DEF.SG.M has  
 trimis fiecare mesager unui aliat pentru a cere ajutor.  
 sent every messenger.ACC a.DAT ally for to ask help  
 ‘Cornered by enemy armies, the emperor has sent every messenger  
 to an ally in order to ask for help.’

**DO<sub>UQ</sub> before IO<sub>ind</sub>, IO clitic**

- b. Încolțit de armatele inamice, împăratul  
 Cornered by armies.DEF.PL enemy.PL, emperor.DEF  
 i-a trimis fiecare mesager unui aliat  
 him.3P.SG.DAT-has sent every messenger.ACC a.DAT ally  
 pentru a cere ajutor.  
 for to ask help  
 ‘Cornered by enemy armies, the emperor has sent every messenger  
 to an ally in order to ask for help.’

**DO<sub>ind</sub> before IO<sub>UQ</sub>, no IO clitic**

- c. Încolțit de armatele inamice, împăratul a  
 Cornered by armies.DEF.PL enemy.PL, emperor.DEF has  
 trimis un mesager fiecărui aliat pentru a cere  
 sent a messenger.ACC every.DAT ally for to ask  
 ajutor.  
 help  
 ‘Cornered by enemy armies, the emperor has sent a messenger to  
 every ally in order to ask for help.’

**DO<sub>ind</sub> before IO<sub>UQ</sub>, IO clitic**

- d. Încolțit de armatele inamice, împăratul  
 Cornered by armies.DEF.PL enemy.PL, emperor.DEF  
 i-a trimis un mesager fiecărui aliat  
 him.3P.SG.DAT-has sent a messenger.ACC every.DAT ally  
 pentru a cere ajutor.  
 for to ask help  
 ‘Cornered by enemy armies, the emperor has sent a messenger to  
 every ally in order to ask for help.’

**IO<sub>UQ</sub> before DO<sub>ind</sub>, no IO clitic**

- e. Încolțit de armatele inamice, împăratul a  
 Cornered by armies.DEF.PL enemy.PL, emperor.DEF has  
 trimis fiecărui aliat un mesager pentru a  
 sent every.DAT ally a messenger.ACC for to  
 cere ajutor.  
 ask help

‘Cornered by enemy armies, the emperor has sent every ally a messenger in order to ask for help.’

**IO<sub>UQ</sub> before DO<sub>ind</sub>, IO clitic**

- f. Încolțit de armatele inamice, împăratul  
 Cornered by army.DEF.PL enemy.PL, emperor.DEF  
 i-a trimis fiecărui aliat un mesager  
 him. 3P.SG.DAT-has sent every.DAT ally a messenger.ACC  
 pentru a cere ajutor.  
 for to ask help

‘Cornered by enemy armies, the emperor has sent every ally a messenger in order to ask for help.’

**IO<sub>ind</sub> before DO<sub>UQ</sub>, no IO clitic**

- g. Încolțit de armatele inamice, împăratul a trimis  
 Cornered by army.DEF.PL enemy.PL, emperor.DEF has sent  
 unui aliat fiecare mesager pentru a cere ajutor.  
 a.DAT ally every messenger.ACC for to ask help

‘Cornered by enemy armies, the emperor has sent an ally every messenger in order to ask for help.’

**IO<sub>ind</sub> before DO<sub>UQ</sub>, IO clitic**

- h. Încolțit de armatele inamice, împăratul  
 Cornered by armies.DEF.PL enemy.PL, emperor.DEF  
 i-a trimis unui aliat fiecare mesager  
 him.3P.SG.DAT-has sent a.DAT ally every messenger. ACC  
 pentru a cere ajutor.  
 for to ask help

‘Cornered by enemy armies, the emperor has sent an ally every messenger in order to ask for help.’

Section 2.2 reports on the experimental results, focusing on the norming tasks. The results regarding the acceptability task will be left for a separate paper due to space limitations.

### 3. Experimental results and discussions

#### 3.1 Experiment 1 – scope dependencies with ditransitives (unmarked DOs)

As already pointed out, Experiment 1 verified scope dependencies between the IO and an unmarked DO. One first observation with respect to the data gathered as part of this experiment is that both arguments seem to allow for both a wide as well as narrow scope reading irrespective of their

relative order with respect to the other, DP type or dative clitic doubling. The surface word order, however, seems to influence scope readings, with the leftmost arguments outscoping the other one in the *IO before DO* configuration. Table 2 shows this at stake:

Nr. crt.	Word order	DP type	+/- IO clitic	Results		
				Narrow scope indefinite DP %	Wide scope indefinite DP %	Both readings %
1.	DO < IO	UQ <sub>DO</sub> & Indef <sub>IO</sub>	no IO cl.	57.63	25.54	16.88
2.	DO < IO	UQ <sub>DO</sub> & Indef <sub>IO</sub>	with IO cl.	38.95	38.80	21.23
3.	DO < IO	Indef <sub>DO</sub> & UQ <sub>IO</sub>	no IO cl.	65.33	16.50	18.15
4.	DO < IO	Indef <sub>DO</sub> & UQ <sub>IO</sub>	with IO cl.	68.71	13.80	17.49
5.	IO < DO	UQ <sub>IO</sub> & Indef <sub>DO</sub>	no IO cl.	75.43	10.93	13.63
6.	IO < DO	UQ <sub>IO</sub> & Indef <sub>DO</sub>	with IO cl.	75.30	11.30	13.40
7.	IO < DO	Indef <sub>IO</sub> & UQ <sub>DO</sub>	no IO cl.	26.17	60.42	13.38
8.	IO < DO	Indef <sub>IO</sub> & UQ <sub>DO</sub>	with IO cl.	18.48	65.50	16.03

**Table 2. Scope dependencies with ditransitives (unmarked DO)**

Let us first consider the *DO before IO* word order:

a) In the *DO<sub>UQ</sub> - IO<sub>Indef</sub>* order most respondents favour a narrow scope interpretation on the indefinite IO (57,63%) and a wide scope one for the DO. The wide scope reading of the indefinite IO is not completely out, however, with 25,54%+16,88% of the respondents allowing for it. The presence of a dative pronominal clitic doubling the IO seems to tilt the balance towards a more prominent wide scope interpretation of this DP (38,80) but the opposite reading remains quite a robust reading variant with 38.95% of the respondents selecting it as the only possible interpretation and 21.23% accepting both the wide and the narrow scope readings.

When we use a *DO<sub>Indef</sub>* in the *DO before IO* order, the respondents seem to prefer a narrow scope interpretation for the DO irrespective of whether there is clitic doubling on the *IO<sub>UQ</sub>* or not. The wide scope reading for the indefinite DO is also available, with a small percentage of respondents

selecting it as the only available interpretation (16.50%/13.80%), while another group consider both readings acceptable (18.15% and 17.49%). It seems then that surface word order does not influence scope readings with the *DO before IO* order.

The DP type seems a stronger parameter in this respect: an unmarked indefinite DO tends to be assigned a narrow scope reading. The same is at stake with the undoubled indefinite IO, which is also preferred with a narrow scope interpretation (57.63%), while with the doubled IO things are not as clear, respondents being divided between a wide scope and a narrow scope reading (38.95% vs. 38.88%). The fact that it is the DP type which seems to favour one reading over another and not word order shows that the configuration itself cannot be blamed for restricting scope interpretation.

Considering now the *IO before DO* surface word order, where we notice the following phenomenon at stake:

a) In the *IO<sub>UQ</sub> before DO<sub>Indef</sub>* surface order, the preferred reading is one in which the indefinite DO has narrow scope both when the *IO<sub>UQ</sub>* has been clitic doubled as well as when there is no dative clitic (75.43% and 75.30% respectively).

b) For the *IO<sub>Indef</sub> before DO<sub>UQ</sub>* surface order, on the other hand, the respondents seem to prefer a wide scope interpretation for the indefinite IO.

Thus, unlike in the previous case (i.e., *DO before IO* surface order) in the *IO before DO* configuration, the wide scope interpretation seems to be assigned to the leftmost DP irrespective of whether this is headed by an indefinite determiner or by a universal quantifier. Importantly, the opposite interpretation remains an option for speakers.

### 3.2 Experiment 2 – scope dependencies with ditransitives (DOMed DOs)

Table 3 shows the scope dependencies from within ditransitive configurations containing a DOMed DO and where all the other parameters have been varied in the same way as in Experiment 1.

In the *DO before IO* order, scope readings seem to change function of the DP type:

a) with the *DO<sub>UQ</sub> before IO<sub>Indef</sub>* configuration, the indefinite IO is preferred on a narrow scope reading irrespective of whether it has been clitic doubled or not. The DO is mostly preferred on a wide scope interpretation. Importantly, however, the opposite reading represents quite a robust option: 20,55% of the respondents assigning the indefinite IO in the *DO<sub>UQ</sub> before IO<sub>Indef</sub>* order a wide

scope reading and 26,83% allowing for both readings on the IO. Similar results are at stake for the clitic doubled IO (27.25% + 20.07%).

b) The  $DO_{Indef}$  before  $IO_{UQ}$  configuration reinforces the idea that the DP type has a say with respect to scope interpretation: the indefinite DOMed DO is preferred on a wide scope reading both when co-occurring with an undoubled IO (50.21%) and when the IO has been clitic doubled (43.68%). A narrow scope interpretation for the DO is also available. Interestingly, when comparing this configuration with the corresponding one in experiment 1, we notice an important difference: the indefinite unmarked DO in Experiment 1 was clearly prone towards a narrow scope reading, while in experiment 2, the speakers tend to assign the DOMed indefinite DO a wide scope reading. Thus, marking has an important impact on scope readings within the same configuration  $DO_{Indef}$  before  $IO_{UQ}$

As noticed from Experiment 1, scope interpretation within the **IO before DO** configuration appeared to be dependent on surface word order, with the IO always getting a wide scope reading on account of being the leftmost DP. The same seems to be at stake with this configuration in Experiment 2. The only difference we noticed with respect to the state of affairs in Experiment 1 is that the differences regarding the interpretation scores are no longer that drastic: the respondents are more flexible in allowing the wide scope reading to the DOMed DO while unmarked DOs in Experiment 1, were more clearly grouped with the narrow scope reading. It seems that DO marking has a say with respect to scope interpretation.

Nr. crt	Word order	DP type	+/- IO clitic	Results		
				Narrow scope indefinite DP%	Wide scope indefinite DP%	Both readings %
1.	DO < IO	UQ <sub>DO</sub> & Indef <sub>IO</sub>	no IO cl.	52.64	20.55	26.83
2.	DO < IO	UQ <sub>DO</sub> & Indef <sub>IO</sub>	with IO cl.	52.67	27.25	20.07
3.	DO < IO	Indef <sub>DO</sub> & UQ <sub>IO</sub>	no IO cl.	38.21	50.21	11.58
4.	DO < IO	Indef <sub>DO</sub> & UQ <sub>IO</sub>	with IO cl.	40.31	43.68	16.01
5.	IO < DO	UQ <sub>IO</sub> & Indef <sub>DO</sub>	no IO cl.	54.09	28.13	17.78
6.	IO < DO	UQ <sub>IO</sub> & Indef <sub>DO</sub>	with IO cl.	50.73	34.06	15.20
7.	IO < DO	Indef <sub>IO</sub> & UQ <sub>DO</sub>	no IO cl.	29.83	52.01	18.71
8.	IO < DO	Indef <sub>IO</sub> & UQ <sub>DO</sub>	with IO cl.	22.01	59.75	18.22

**Table 3. Scope dependencies with ditransitives (DOMed DO)**

As a general conclusion for the *DO before IO*, the DOMed DO seems to favour a wide scope reading irrespective of whether it is headed by an indefinite determiner or by a universal quantifier. In the *IO before DO*, the surface word order seems to have precedence over other parameters forcing a wide scope interpretation on the IO which is the leftmost DP (just like in Experiment 1). DOM marking of the DO influences interpretation to a certain extent: respondents seem more prone to allow for a wide scope reading on the DO in this case than they seemed to be with respect to unmarked DOs.

### 3.3 Experiment 3 – scope dependencies with ditransitives (CDed+DOMed DOs)

Just like in Experiment 2, the *DO before IO* order is sensitive to the DP type: CDed+DOMed DOs are usually interpreted on a wide scope reading irrespective of whether they are headed by an indefinite determiner or by a universal quantifier:

a) *DO<sub>UQ</sub> before IO<sub>Indef</sub>*: as shown in Table 4, the indefinite IO acquires a narrow scope reading both when accompanied by a dative clitic and when surfacing by itself (54.85% and 59.52% respectively). The wide scope reading on the IO remains an option for respondents however (17.83%+22.61 and 27.15%+17.97%). Conversely, the CDed+DOMed DO mostly acquires a wide scope reading within this configuration; the narrow scope interpretation is also available.

b) *DO<sub>Indef</sub> before IO<sub>UQ</sub>*: there is a clear tendency from the part of the respondents to assign a wide scope reading on the indefinite CDed+DOMed DO. The narrow scope interpretation remains an option though and, interestingly, when the IO has been clitic doubled, the respondents seem to assign a wide scope reading to it more readily, which leads to a uniformization regarding scores (40.76% vs. 44.72%). The same effect may be observed in Experiment 2 when a clitic doubled IO co-occurs with a DOMed DO (40.31% vs. 43.68%).

In the previous two experiments the **IO before DO** order showed a clear preference for a wide scope interpretation on the leftmost DP. The presence of a CDed+DOMed DO seems to change matters to a certain extent:

a) *IO<sub>UQ</sub> before DO<sub>Indef</sub>*: unlike in the previous two experiments where respondents strongly favoured a narrow scope reading for the DO, the CDed+DOMed DO is preferred on a wide scope reading now. This points to the fact that DP type takes precedence on configuration.

b) *IO<sub>Indef</sub> before DO<sub>UQ</sub>*: in this configuration the indefinite IO was assigned a wide scope interpretation by most respondents, being the leftmost DP. Things seem to change, however, when the DO is clitic doubled and differentially marked as participants tend to assign a narrow scope reading on the IO this time (45.64) and to thus favour a wide scope interpretation on the DO. Even more interestingly, the scope preferences of the speakers reverse when the IO has been clitic doubled: the IO is now preferred on a wide scope reading again (42.82%). The narrow scope interpretation remains a very robust option, nevertheless.

Nr · crt ·	Word order	DP type	+/- clitic IO	Results		
				Narro w scope indefi nite DP%	Wide scope indefi nite DP%	Both readi ngs %
1.	DO < IO	UQ <sub>DO</sub> & Indef <sub>IO</sub>	no IO cl.	59.52	17.83	22.61
2.	DO < IO	UQ <sub>DO</sub> & Indef <sub>IO</sub>	with IO cl.	54.85	27.15	17.97
3.	DO < IO	Indef <sub>DO</sub> & UQ <sub>IO</sub>	no IO cl.	27.06	62.27	10.65
4.	DO < IO	Indef <sub>DO</sub> & UQ <sub>IO</sub>	with IO cl.	40.76	44.72	14.50
5.	IO < DO	UQ <sub>IO</sub> & Indef <sub>DO</sub>	no IO cl.	36.61	51.50	11.87
6.	IO < DO	UQ <sub>IO</sub> & Indef <sub>DO</sub>	with IO cl.	36.78	50.86	12.36
7.	IO < DO	Indef <sub>IO</sub> & UQ <sub>DO</sub>	no IO cl.	45.64	28.59	25.75
8.	IO < DO	Indef <sub>IO</sub> & UQ <sub>DO</sub>	with IO cl.	38.48	42.82	18.65

**Table 4. Scope dependencies with ditransitives (CDed+DOMed DO)**

### 3.4 Joining threads together

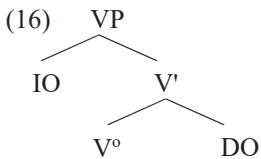
Though scope dependencies were shown to vary in the three experiments function of the DP type and word order, one thing is clear: no configuration is restricted to one scope interpretation only, which shows that Romanian ditransitives do not exhibit scope freezing effects as it is the case for Russian, Japanese, and English.

Nevertheless, some preferences regarding scope interpretation have been uncovered, which might tell us something about the underlying configuration with ditransitives:

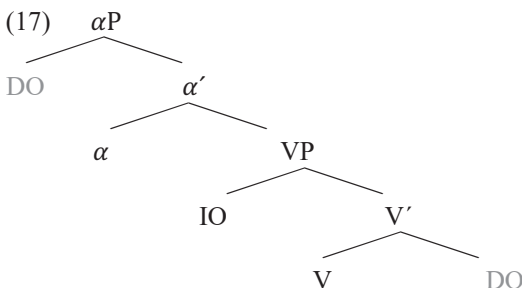
a) **IO before DO**: Word order seems to be quite important in this particular case, influencing scope readings in such a way that the leftmost DP gets a wide scope interpretation (as shown in Experiment 1 & 2). Note, however,

that the opposite reading is never completely discarded and that it becomes quite a robust option with DOMed DOs and even more so with CDed+DOMed DOs, which seem to be generally preferred on a wide scope interpretation.

Under these circumstances, we cannot speak about scope freezing effects with the *IO before DO* configuration. We may, however, conclude that this order represents the basic order of the two internal arguments: according to (16), the IO is merged above DO, which allows us to expect a narrow scope reading on DO.



Starting from such a basic configuration also enables us to explain why the opposite *DO before IO* order seems to be more flexible with respect to scope interpretation: this configuration is a derived one, obtained by having moved the DO to a position above the one occupied by the IO. As a consequence, the DO may be interpreted either in its landing site or reconstructed in its merge position: when the former situation obtains, the DO acquires a wide scope reading, while interpreting it in its merge position allows the DO to have a narrow scope reading.



Finally, one needs to also account for the fact that the strong bias towards interpreting the leftmost DP on a wide scope reading uncovered in the first experiment featuring unmarked DOs, seems to be corrected to the contrary when the DO bears some marking – this is not so spectacular with merely DOMed DOs (though there is a clear influence) but it is clearly



visible with CDed+DOMed DOs. One way to account for this is to imagine that marked DOs need to leave their merging site for feature checking reasons and thereby ending up in positions wherefrom they c-command the IO. (see Tigău 2020 who proposes such an account by building on López 2012).

b) **DO before IO**: as already pointed out, this order allows for more flexibility with respect to scope interpretation on account of it hosting two DO positions which may be interpreted alternatively in relation to the IO. Furthermore, just like in the *IO before DO* configuration, DO marking seems to influence scope readings, respondents preferring a wide scope interpretation for DOMed DOs and CDed+DOMed DOs.

## 4. Conclusions

Unlike English, Japanese, or Russian, where we find scope freezing effects with at least one of the surface word orders, Romanian allows for wide and narrow scope interpretations with both internal arguments irrespective of their ordering one with respect to the other.

Nevertheless, the *IO before DO* configuration seemed to favour interpretations where the leftmost DP received a wide scope reading. Such interpretive effects are taken as evidence that the basic underlying configuration for Romanian ditransitives is the *Goal-over-Theme* one. This account also explains why scope readings within the *DO before IO* surface word order seem to be more flexible: this is a derived configuration which obtains by scrambling the DO to a landing site above the IO. This, in turn, renders available two positions where one may interpret the DO, with the corresponding consequences on scope reading: the DO landing site (wide scope reading) and the DO merge position (narrow scope reading).

The DO type employed i.e., whether it is unmarked vs. DOMed or CDed+DOMed seems to also have a say in the matter in that participants tend to assign a wide scope reading to the marked DOs irrespective of surface word order. It was argued that these DPs exit their merge position from within the VP reaching a landing site which is hierarchically superior to the position hosting the IO.

## References

- Antonyuk, Svitlana. 2015. “Quantifier Scopepe and Scope Freezing in Russian”. PhD diss., Stony Brook University.
- Aoun, Joseph and Audrey Yen-Hui Li. 1989. “Scope and constituency”. *Linguistic Inquiry* 20 (2): 141–172.
- Barss, Andrew and Howard Lasnik. 1986. “A note on anaphora and double objects”. *Linguistic Inquiry* 17 (2): 347–35.
- Cornilescu, Alexandra, Anca Dinu and Alina Tigău. 2017a. “Experimental Data on Romanian Double Object Constructions”. *Revue Roumaine de Linguistique* LXII (2): 157–177.
- Cornilescu, Alexandra, Anca Dinu and Alina Tigău. 2017b. “Romanian Dative Configurations: Ditransitive Verbs, A Tentative Analysis”. *Revue Roumaine de Linguistique*, special volume *Investigating Romanian Datives / Recherches autour du datif roumain* LXII (2): 179–206.
- Hoji, Hajime. 1985. “Logical Form constraints and configurational structures in Japanese”. PhD diss., University of Washington, Seattle.
- Larson, Richard Kurth. 1990. “Double objects revisited: Reply to Jackendoff”. *Linguistic Inquiry* 21 (4): 589–632.
- López, Luis. 2012. *Indefinite Objects. Scrambling, Choice Functions, and Differential Marking*. Cambridge, Massachusetts: MIT
- Marantz, Alec. 1993. “Implications of asymmetries in double object constructions.” In *Theoretical Aspects of Bantu Grammar* edited by Sam A. Mchombo, 113–150. Stanford, CA: CSLI Publications (= CSLI lecture notes 38).
- Miyagawa, Shigeru and Takae Tsujioka. 2004. “Argument structure and ditransitive verbs in Japanese”. *Journal of East Asian Linguistics* 13 (1): 1–38.
- Tigău, Alina. 2020. *Experimental insights into the syntax of Romanian ditransitives*. Berlin/Boston: Walter de Gruyter.

## Notes

---

<sup>1</sup> I would like to thank the audience of the conference AICED 2019 in Bucharest for their critical comments and an anonymous reviewer for helpful comments. The research for this paper has been funded by the Alexander von Humboldt Foundation.

<sup>2</sup> Antonyuk (2015) distinguishes between three classes of ditransitives function of their scope behavior:

Class 1: DP<sub>ACC</sub> > DP<sub>OBL</sub> (ambiguous)  
 DP<sub>OBL</sub> > DP<sub>ACC</sub> (frozen)

Class 2:  $DP_{OBL} > DP_{ACC}$  (ambiguous)  
 $DP_{ACC} > DP_{OBL}$  (frozen)

Class 3:  $DP_{OBL} > DP_{ACC}$  (ambiguous)  
 $DP_{ACC} > DP_{OBL}$  (ambiguous)

## CHAPTER THREE

# THE I\* SINGLE ARGUMENT INTRODUCER: A SOLUTION FOR REPRESENTING THE *BENEFICIARY* ARGUMENT OF CREATION VERBS IN BRAZILIAN PORTUGUESE

ANA REGINA CALINDRO

This paper aims to propose a representation for the *beneficiary* argument of ditransitive sentences with creation verbs (*bake, build, paint, prepare*) in Brazilian Portuguese (BP), as (1): ‘*Maria preparou o jantar para o João/para ele*’ (lit. Mary prepared the dinner to the João/to him). The *beneficiary* argument in sentences such as (1) is introduced by the preposition *a* in European Portuguese (EP), but with *para* in BP. Even though EP and BP shared a common variety in the past, since the 18<sup>th</sup> century, the latter has undergone several diachronic changes. One of them is the substitution of *a* to *para* to denote the *beneficiary* (Torres Morais and Berlinck 2018). In this paper, I will show diachronic data from Calindro (2015 a, b) from the 20<sup>th</sup> century to confirm this substitution and discuss the consequences in terms of argument structure in BP. Contrasting BP data with works on English and Romance languages – Spanish, Romanian and EP (cf. Pylkkänen 2002, Cuervo 2003, Diaconescu and Rivero 2007), I argue that applicative heads are not available for BP ditransitives, because the aforementioned diachronic change attests that all prepositions in BP are transitive elements (cf. Cuervo 2010, Svenonius 2004). Hence, a prepositional head *p* is perfectly capable of introducing the IA in the argument structure of BP (cf. Calindro 2016, 2020; Svenonius 2003, 2004; Wood 2012). However, the representation with pP does not account for the two semantic readings *para* conveys for the *beneficiary* in (1). In this example, the DP *o João* is either *beneficiary of the theme* – dinner, or only *beneficiary of the event* of *Maria* having prepared dinner, as it would be if introduced by *por* (for) as well. I

will thus show that the representation of these sentences with *i\** single argument introducer (cf. Wood and Marantz 2017) provides a better solution for these two semantic readings.

Keywords: creation verbs; beneficiary; argument structure; diachronic change; Brazilian Portuguese

## 1. Introduction<sup>1</sup>

The aim of this paper is to propose a representation of the argument structure of ditransitive sentences with creation verbs in Brazilian Portuguese (BP):

- (1) Maria preparou o jantar para o João / para ele. (BP)  
 Maria prepared DEF dinner P<sub>para(to)</sub> DEF João.OBL/ for him.3SG  
 ‘Maria prepared dinner for João / for him.’

Example (1) shows that preposition *para* introduces the *beneficiary* indirect argument (IA). In European Portuguese (EP) (cf.2), on the other hand, preposition *a* is used in the same context. Thus, the data from modern BP and EP indicates that the former has undergone a diachronic change regarding the introduction of *beneficiary* IAs.

- (2) A Maria preparou o jantar ao João /  
 DEF Maria prepared DEF dinner P<sub>a(to)</sub> the João.DAT /  
 preparou-lhe o jantar<sup>2</sup>. (EP/ \*BP)  
 prepare – CL.3P.SG.DAT DEF dinner  
 ‘Maria prepared dinner for João / for him.’

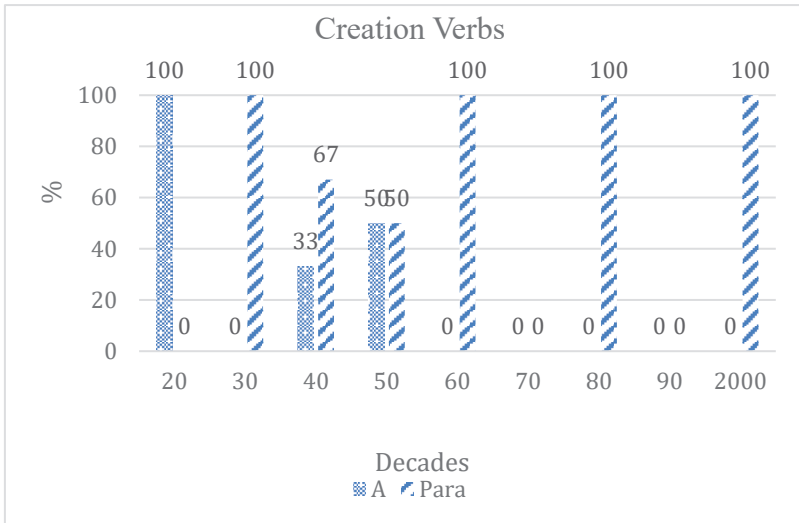
Historically, BP and EP used to share the same argument structure. Hence, only preposition *a* was used to introduce the *beneficiary* with creation verbs. However, several diachronic works have observed several changes in BP, since the 18<sup>th</sup> century, one of them related to the prepositions that introduce arguments (Galves 2001, Kato et al. 2009). Specifically, in the 19<sup>th</sup> century, preposition *para* started to substitute *a*, and eventually replaced the latter completely in sentences such as (1). Additionally, the IA in BP cannot alternate with third person dative clitics anymore (*lhe/lhes*)<sup>3</sup>. For instance, in (1), the IA alternates with a full pronoun preceded by a lexical preposition (cf. Torres Morais 2007, Torres Morais and Salles 2010; Torres Morais and Berlinck 2007, 2018, Calindro 2015a, 2015b, 2016).

According to the literature, creation verbs entail the making of an entity denoted by the direct object<sup>4</sup>, which can be concrete (cf.5) or abstract (cf.6) (cf. Levin 1993, Amaral and Cançado 2014):

- (3) a. Clarice Lispector wrote a book.
  - b. Tarsila do Amaral ainted Abapuru.
  - c. They built a house.
  - d. Paola baked a cake.
  
  - (4) a. Adriana Calcanhoto composed a song.
  - b. Hypatia formulated a mathematical theorem.
- (adapted from Amaral and Cançado 2014, 52)

Calindro (2015a) searched on a corpus from the 20<sup>th</sup> century – the period when the main changes regarding prepositions heading IAs were established in BP, including sentences with creation verbs. The data analyzed were collected from a book (entitled *Primeira página: 90 anos de história nas capas mais importantes da Folha – 90 years of history through the most important covers of Folha*), which comprises 223 covers from *Folha de São Paulo* – a major Brazilian newspaper – that spans the 20<sup>th</sup> century from 1920 to 2010<sup>5</sup>, allowing a thorough overview of the 20<sup>th</sup> century.

The search for ditransitive sentences in the corpus resulted in a total of 592 sentences with verbs of transfer (*give, send*), movement (*throw*), and creation verbs accompanied by IAs introduced by *a* or *para*. From this total, only 18 sentences containing creation verbs were found. The small number of ditransitives featuring creation verbs may be accounted for by the inappropriateness of the material analyzed, as newspapers are not the ideal source to extract sentences with creation verbs. However, even though this result was not expressive, it sheds light on the argument structure of creation verbs, mainly when synchronic data from both BP and EP are compared, i.e., in modern BP preposition *a* was completely replaced by *para*<sup>6</sup>. The results showed that preposition *para* completely replaced *a* with creation verbs during the 20<sup>th</sup> century, mainly from the 60s onwards:



**Graph 1. Creation verbs results throughout decades**

The graph illustrates that prepositions *para* and *a* were in complementary distribution at the beginning of the century, as the examples below show.

- (5) Uma provocação a mais á policia e o continuar  
 One provocation more to police.DEF and DEF continuity  
 a escrever cartas aos jornaes.  
 to writing letters P<sub>a</sub>(to).DEF newspapers  
 ‘One more provocation to the police and (they) continued writing  
 letters to the newspaper  
 (05/06/1926)<sup>7</sup>
- (6) (...) ambos os paizes regulavam todas as questões  
 (...) both DEF countries regulated all DEF questions  
 que havia surgido com a dissolução do  
 Estado polonez,  
 that there.had appeared with DEF dissolution of.DEF State  
 polish  
creando dessa maneira uma base solida para a paz européa.  
 creating of-this way a base solid P<sub>para</sub>(to) DEF peace European

‘(...) both countries regulated all questions that had appeared with the dissolution of the Polish State, thus creating a solid foundation for European peace’.  
(30/09/1939)

As the data show, there was a significant change in the 60s, when preposition *para* completely substituted preposition *a*. Additionally, as only 18 examples with creation verbs were found in the corpus, besides the percentage, it is important to show their exact distribution in order to support this discussion<sup>8</sup>. Thereby, I present below a table with the results per decade:

Decades	20	30	40	50	60	70	80	90	2000
<i>a</i>	1	0	1	2	0	0	0	0	0
<i>para</i>	0	1	2	2	3	0	4	0	2

**Table 1. Distribution of *para* and *a* with creation verbs per decade.**

Besides sentences with creation verbs, in contexts with movement and transfer verbs, the IA introduced by *a* co-occurs with *para*. Hence, preposition *a* is grammatical in BP (cf. Calindro 2015a, 2020) – ‘*Maria enviou uma carta ao/para o João*’ (Maria sent a letter to John). Sentence (2), however, is ungrammatical for Brazilian speakers.

Furthermore, with creation verbs, *para* conveys two semantic readings for the *beneficiary* IA. In (1), the DP *o João* is either the *beneficiary of the theme* – dinner, or only *beneficiary of the event* of *Maria* having prepared dinner, as it will be discussed further in this text.

Given the discussion above, there are two questions to address: *i.* what is the syntactic representation of these structures from BP? *ii.* Is it possible to syntactically represent the two semantic readings for the IA introduced by *para* in BP?

Therefore, in this paper, I will expand the discussion in Calindro (2015 a, b et seq.), in order to answer these questions. In Section 2, I will explore the discussion for creation verbs in English, Bantu languages (specifically Chichewa and Chaga) (Larson 1988, Marantz 1993, Pytkänen 2002), Spanish (Cuervo 2003), Romanian (Diaconescu and Rivero 2007), and European Portuguese (Torres Morais 2007). Subsequently, I compare these languages to the data from BP, analyzing the status of the prepositions that introduce IAs. I conclude that applicative heads are not part of the argument structure of BP and the IAs can be introduced by prepositional phrases (pP).



In Section 3, the theory developed by Wood and Marantz (2017) will be explored to explain these types of arguments using a system that can account for the semantic differences conveyed by *para*. Wood and Marantz's proposal provides an effective representation of the argument structures at hand. Additionally, I will show that BP data can perfectly illustrate the facts they discuss with more elements, as we will see in Section 3, and conclude the paper in the final remarks on Section 4.

## 2. The syntactic representation of ditransitive structures with creation verbs in BP

### 2.1 Ditransitive sentences - DOCs and PDCs

The literature provides two main solutions for events with two arguments. Following the seminal works of Baker (1988) and Larson (1988), the discussion regarding these structures was based on the following sentences in English:

- (7) John gave a cake to Mary.  
 (8) John gave Mary a cake.

These two sentences are part of the phenomenon known as *dative alternation*. There are morphological and semantic aspects that restrict the *dative alternation* in English. According to Mazurkewich (1984), verbs with a Germanic root (*tell, show, get*) allow both constructions whereas verbs with a Latin root (*report, demonstrate, obtain* – cf. 9) only allow Prepositional Dative Constructions (PDC) as in (7)<sup>9</sup>.

- (9) a. Peter told / \*reported his boss the news.  
 b. The professor showed / \*demonstrated us his new methods.  
 c. Paul got / \*obtained his girlfriend a ticket.  
 (Mazurkewich 1984, 83)

Semantically, Double Object Constructions (DOCs – e.g. (8)) and PDCs convey different interpretations (cf. Larson 1988, Pinker 1989). For instance, the dative object of a DOC has to be a viable [+animated] possessor, and the event should entail a transfer of possession between the *theme* (DO) and the *goal* (IO). Additionally, some verbs allow DOCs and others do not, due to the semantic content of the preposition in the

construction. In (7), preposition *to* is a dative Case assigner without semantic content. When the preposition does have semantic content, *dative alternation* is not possible, as with oblique instrumentals (10) and locative phrases (11):

- (10) I cut the salami with a knife.  
 \* I cut a knife the salami.
- (11) John left his books on the sofa.  
 \* John left the sofa his books.

(Larson 1988, 372)

Larson (1988) proposes a derivational approach to represent DOCs and PDCs, according to which the DOC would derive from the PDC in a VP-Shell representation, where the VP would have layers to support both arguments. Although English presents several irregularities regarding *dative alternation*, Larson points out that, in languages with morphological *applicatives*, as Bantu languages, DOCs and PDCs alternate in a more productive way.

Hence, following Larson's comment on Bantu languages, Marantz (1993) introduced applicative heads as a solution for representing DOCs in English. The author compares the following structures:

#### *English*

- (12) a. I baked a cake.  
 b. I baked him a cake.  
 c. I ran.  
 d. \*I ran him.

#### *Chichewa* (Alsina and Mchombo 1993)

- (13) Chitsiru chi-na-gul-ir-a                      atsikana mpmatso  
 Fool      SP-pst-buy- APPL-fv<sup>10</sup>      girls      gift  
 'The fool bought a gift for the girls'

#### *Chaga* (Bresnan and Moshi 1993)

- (14) a. N- à - à - lyì - í - à                                      m-kà      k-élyá.  
 FOC - 1P.SG - PRS- eat - APPL - fv                      I-wife      7-food  
 'He is eating food for his wife'

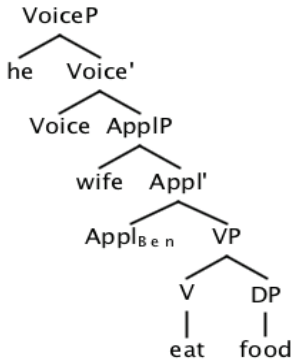
- b. N- à - i - zric - í - à                      *mbùyà.*  
 FOC - 1P.SG - PRS - eat - APPL – fv        *9-friend*  
 ‘He is running for a *friend*’  
 (Marantz 1993, 121)

These examples are particularly interesting for the discussion in this paper, because the main example in English displays a creation verb (*bake*), not with a transfer verb (*give*), as most examples in Larson’s paper. The IA in (12) ‘him’ and the IA in (13) ‘girls’ are both the *beneficiary* of the Themes – ‘cake’ and ‘gift’, respectively. Therefore, the proposal in Marantz’s paper may fit well for the data in BP I am analyzing here.

Differently from English (12), Chichewa (13), and Chaga (14) add an applicative morpheme to denote the *beneficiary*. Therefore, based on the applicative morphology, Marantz proposed an applicative analysis that differs from Larson’s derivational one. Marantz assumes DOCs have an applicative head to introduce the IA from (12). According to the author, the applicative head takes an event predicate (a VP) – and introduces an argument, which is thematically related to the event described by the verb – that is why ApplP can be used in ditransitives. In PDCs, Marantz proposes the IA is a PP containing a lexical preposition.

Following Marantz (1993), but focusing on the semantics of the events, Pylkkänen (2002) points out that the similarities among these benefactive structures in English and Bantu languages are only apparent. For instance, in example (14) from Chaga, the applicative head establishes a relation between the *beneficiary* and the event described by the VP whereas in English (12) the ApplP relates the Theme (DO) to the *beneficiary*. Thus, the author proposes two types of applicatives, naming them *high* and *low applicatives*.

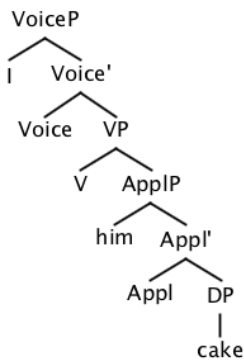
The high applicative is very similar to Voice because it adds a new participant to the event described by the verb, being subcategorized by it. In (14), ‘he’ ate the food instead of ‘his wife’, i.e., executed the event in her place (for her). Hence, ApplP is higher than the VP, thus *high applicative* (15).

(15) High Applicative (Chaga)<sup>11</sup>

(Pylkkänen 2002, 19)

The interpretation represented in (15) – which shows that the agent ate the food because the *beneficiary* (wife) could not do it, or did not want to do it – is not possible in English. The only feasible interpretation in (16) is that the ‘cake’ was baked for ‘him’ to eat. Hence, there is an implicit intention that he will be the *beneficiary* of the cake, the theme of the event. According to Pylkkänen (2002), this relation may be expressed by a *low applicative* head. Thus, DOCs in English have low applicatives, as illustrated below:

## (16) Low Applicative (English)



(Pylkkänen 2002, 19)

In BP, however, both low and high applicative semantic readings are possible when the *beneficiary* with a creation verb is introduced by *para*.

Consequently, in the next sections, I will discuss if the applicative approach suits the BP data.

## 2.2 Ditransitive sentences in Romance languages

As we will see further in this paper, these seminal discussions for English presented in the previous section were crucial for understanding how structures with creation verbs behave across languages. Therefore, many authors proposed that languages such as Spanish, Romanian, EP, and BP have applicative heads as well. In this article, I will argue that the works, which propose an applicative structure for BP are not considering several characteristics that set BP apart from English and other Romance languages (cf. Calindro 2020).

Starting with an overview of Romance languages, Cuervo (2003) for Spanish and Diaconescu and Rivero (2007) for Romanian propose *dative alternation* and applicative heads in these languages have different characteristics from the ones observed in English. For instance, clitic doubling is obligatory in Spanish. In Spanish, clitic (*le*) co-occurs with the DP introduced by a dummy preposition (cf. 17a). In Romanian, the goal ‘Mariei’ marked with dative case may be doubled by a dative clitic (*îi*) (cf. 19a)<sup>12</sup>. Hence, the clitic is the Spell-out of the applicative projection, because it is responsible for lexicalizing the number and person features of the DP which is in SpecAppIP (17b) and (19b).

### *Spanish*

- (17) a. Pablo le mandó un diccionario  
 Pablo CL.3P.SG.DAT.APPL sent a dictionary  
 a Gabi.  
 to Gabi.DAT  
 ‘Pablo sent Gabi a dictionary’.
- b. [<sub>VoiceP</sub> Pablo [<sub>v'</sub> voice [<sub>VP</sub> mandó [<sub>AppIP</sub> a Gabi [<sub>APPL</sub> le [<sub>DP</sub> un diccionario]]]]]]
- (18) Pablo mandó un diccionario a Gabi/ a Barcelona.  
 Pablo sent a dictionary to Gabi.OBL/ to Barcelona. OBL  
 ‘Pablo sent a dictionary to Gabi/to Barcelona’.

(Cuervo, 2003, 51)

*Romanian*

- (19) a. Mihaela îi trimite Mariei o scrisoare.  
 Mihaela CL.3P.SG.DAT sent Mary.DAT a letter  
 ‘Mihaela sends Mary a letter’.
- b. [<sub>VoiceP</sub> Mihaela [<sub>V'</sub> voice [<sub>VP</sub> trimite [<sub>AppIP</sub> Mariei [<sub>APPL'</sub> îi [<sub>DP</sub> o scrisoare]]]]]]]
- (20) Mihaela trimite Mariei o scrisoare.  
 Mihaela sent Mary.DAT a letter  
 ‘Mihaela sends a letter to Mary’.  
 (Diaconescu and Rivero 2007, 21,23)

Therefore, (17) and (19) are applicative constructions, and alternate with constructions with different syntactic structures as in (18) and (20)<sup>13</sup>.

EP, according to Torres Morais (2007), also has applicative constructions, even though there is no obligatory clitic doubling in this language. In sentence (21), preposition *a* is a dummy element that lexicalizes the applicative head. The main characteristic for this assumption is that the IA can always alternate with the dative clitic *lhe(s)*. Hence, even though clitic doubling is not obligatory in EP, the clitic can only be used when the IA is introduced by *a*, and it cannot alternate when introduced with *para*. This fact gave the grounds for the author to propose (21) is a DOC and (22) a PDC.

- (21) a. O José forneceu/ofereceu ajuda aos pobres/  
 DEF José provided/offered help P<sub>a(to)</sub>. DEF poor. DAT /  
ofereceu-lhes ajuda.  
 offered-3P.PL.DAT help  
 ‘José offered the poor help / offered them help’.
- b. [<sub>VP</sub> O João [<sub>V'</sub> V [<sub>VP</sub> ofereceu [<sub>AppIP</sub> aos pobres/lhes [<sub>APPL'</sub> Ø [<sub>DP</sub> ajuda]]]]]]]
- (22) O José forneceu/ofereceu ajuda para os pobres através  
 DEF José provided/offered help P<sub>para(to)</sub> DEF poor.OBL through  
 da ONU/ \*ofereceu-lhes ajuda.  
 of.DEF UN / \* offered-CL.3P.PL.DAT help  
 ‘José offered the poor help through the UN/ offered them help’.  
 (adapted from Torres Morais 2007, 101)

Furthermore, in (21), there is a direct transfer between the DO and the IO, as attested for DOCs in English. In sentence (22), on the contrary, there is an indirect transfer, emphasized by ‘through the UN’, the IO is introduced by *para*, and it cannot alternate with dative clitics, hence it is a PDC. In the DOC (21), EP has inherent Case and the preposition *a* introduces the IA associated with the  $\theta$ -role related to them, which is *goal* in this example. Additionally, the inherent dative must correspond to the morphological case, in EP, thus the dative clitic is the morphological expression of the applicative head, introduced as a proper argument in SpecAppIP (cf. 21b) (cf. Torres Morais and Salles 2010). In the PDC (cf. 22), the IA is introduced by a PP.

Additionally, *dative alternation* in EP would only hold if we consider that EP presents two types of ditransitive sentences because in one set of examples it is possible to alternate the IA with the dative clitic (cf. 21), and in the other, it is not (cf. 22). Hence, the status of preposition *a* is different from *para* - the former being a functional element and the latter a lexical one.

In modern EP, however, *a* seems to be allowed in contexts that were exclusive of *para*, with locatives, for example. Thus, preposition *a* may have become the *superset* of prepositions in EP like *para* is in BP, insofar as *para* diachronically became the preposition which denotes *goals, locatives, and beneficiaries*. In EP, preposition *a* appears to be going in the same direction.

According to the literature (Brito 2009; Torres Morais and Salles 2010) the IO locative is introduced by *para* and it cannot alternate with the dative clitic *lhe(s)*:

- (23) A Maria enviou (\**lhe*) uma carta para Lisboa.  
 DEF Maria sent (CL.3P.SG.DAT) a letter P<sub>para(to)</sub> Lisbon.OBL  
 ‘Maria sent a letter to Lisbon’.

However, when searching for examples for this paper, I found the following example on a website about Portuguese grammar (Flip.pt) with preposition *a* introducing a pure locative:

- (24) DEF Manuel já chegou a Lisboa.  
 The Manuel already arrived P<sub>a(to)</sub> Lisbon.  
 ‘Manuel has already arrived in Lisbon’.

(FLiP.pt *n.d.*)

Hence, preposition *a* occurs in contexts where only *para* used to be possible in EP. The details related to this assumption I will leave for future work.

BP, on the other hand, reanalyzed the introduction of IAs to the opposite direction. In some contexts, preposition *para* substituted *a*, and in others they co-occur. BP seems to have only Structural Case, because preposition *a* and *para* cannot alternate with the morphological 3<sup>rd</sup> person counterpart *lhe(s)*. Hence, BP presents only prepositional constructions, as I will discuss in the next section.

### 2.3 Brazilian Portuguese

Considering ditransitive sentences with transfer verbs (*send*), as (25), Armelin (2011) assumes applicative constructions for BP, mostly based on a specific dialect from a region in *Minas Gerais*<sup>14</sup> – a state from the southwest of Brazil, that allows the preposition which heads the IA to be suppressed (cf. 26 and 27):

- (25) Maria enviou uma carta    ao/para    o João / ele.  
 Maria sent    a    letter    P<sub>a/para (to)</sub> DEF João / to him. 3P.SG  
 ‘Maria sent a letter to John’.
- (26) O João deu o livro    ao/para o Pedro.  
 DEF João    gave DEF book    P<sub>a/para (to)</sub> DEF Pedro  
 ‘John gave the book to Pedro’.
- (27) O João deu o livro o Pedro.  
 DEF João    gave DEF book    DEF Pedro  
 ‘John gave Pedro the book’.

According to the author, based on data collected by Scher (1996), the suppression of the preposition which introduces the IA is only possible when preposition *a* is one of the options (cf. 26), and the IA has the  $\theta$ -role of *goal*. Hence, *a* would be a Case assigner and the true dative argument introducer in these types of constructions. Therefore, following Pujalte (2010), the IAs introduced by *a* are true dative arguments, because even when they are omitted, it is possible to infer their presence. Thus, this *a* would be part of the IA of true ditransitive sentences which can be introduced in the argument structure by the functional projection ApplP.



On the other hand, the IAs in sentences with creation verbs cannot be omitted or inferred (cf. 28). Furthermore, they cannot be introduced by *a* in BP, only *para*, as mentioned previously. Additionally, the *beneficiary* of an event can be added to any verb and any type of construction (cf. 29).

- (28) a. O João construiu a casa \*ao / para o Pedro.  
 DEF João built a house P *a/para (to)* DEF Pedro
- b. \*O João construiu a casa o Pedro.  
 DEF João built DEF house DEF Pedro  
 ‘John built Pedro a house’
- (29) A Maria entregou o livro ao /para o Pedro  
 DEF Maria gave DEF book P *a/para (to)* DEF Pedro  
 \*a/ para o João.  
 P*a/para (to)* DEF João  
 ‘Maria gave the book to Pedro for João’

In (29), ‘ao/para o Pedro’ is the *goal*, as it can be introduced by both *a* and *para*, hence it could be represented with an applicative head. According to Armelin (2011), as the *beneficiary* ‘\*a/para o João’ cannot be introduced by *a*, only *para* in BP, it is not introduced by an ApplP, but via an adjunction. I assume there are issues with this analysis, as I will discuss further in this paper.

As we saw in the previous section, in the *dative alternation* examples there are two arguments related to the event. In the argument structure, the DO is licensed by the verb as its selection depends on the lexical content of the verbal root<sup>15</sup>. Subjects, as well as indirect arguments, are not licensed by the VP – as Larson (1988) proposes – because they are extra participants of the event and both can be omitted without compromising meaning. Therefore, they have structural meaning, thus they should be licensed syntactically and semantically by specialized heads such as the ApplP, as proposed by Marantz (1993) and Pylkkänen (2002). Notice, these authors do not distinguish the  $\theta$ -roles of the IAs, i.e., the same head introduces *goals* and *beneficiaries*.

Therefore, differently from Armelin’s proposal for BP, the works of Larson (1988), Marantz (1993), and Pylkkänen (2002) do not distinguish *goals* from *beneficiaries*. Furthermore, Marantz’s arguments for proposing ApplP for English are data from Bantu languages (see Section 2.1) whose *beneficiaries* have applicative morphology. Additionally, his main example

in English is with a creation verb ('bake' – cf. 12) with a *beneficiary* argument. Hence, this discussion points out for the possibility to analyze *goals* and *beneficiaries* as the same category, not the former being an ApplP and the latter an adjunct phrase, as Armelin proposes.

On the other hand, according to Pujalte, the separation among *goals*, *beneficiaries* and *locatives* seems to be more evident in Romance languages. Guéron (1985) assumes that preposition *a*, in Romance languages, like Spanish, is related to the dative case and the *goal*  $\theta$ -role, which is a primary  $\theta$ -role. Additionally, preposition *a* may be associated with secondary  $\theta$ -roles *beneficiary* and *locative*<sup>16</sup>. The same is true for EP. In BP, on the other hand, preposition *a* behaves differently from other Romance languages, because it comprises *goal* and *locative*  $\theta$ -roles. Additionally, *locatives* can be introduced by *para*, but the *beneficiary* is only introduced by *para* in BP, never by *a*.

Therefore, the main reason for assuming applicative heads in BP does not hold, as it is based on the allegedly different status prepositions *a* and *para* may have when they introduce *goals* in the argument structure, when they actually have the same status. Hence, in the next section, I will discuss that the diachronic change described regarding the introduction of IAs reaffirms prepositions *a* and *para* in BP are different from the same pair in EP.

### 2.3.1 Prepositions in BP

In this section, I will explore the specific characteristics of prepositions in ditransitive sentences in BP. As mentioned before, preposition *a*, preferably used in EP, has been gradually replaced by other prepositions in BP such as *para* (to), *de* (of), *em* (in), *com* (with) in several contexts, since the 19<sup>th</sup> century (cf. Kato et al. 2009). Differently from EP, these prepositions in BP are not functional elements, as discussed in Calindro (2015a, 2015b), they are lexical prepositions that assign oblique Case to their complement. As exemplified in (21) and (22) for EP, the IA in (21) can alternate with dative clitics because *a* is a functional element responsible for assigning dative Case to it. In (22), however, the IA is introduced by *para* and thus cannot alternate with dative clitics. Therefore, *para* is a lexical preposition.

A lot has been discussed in the literature regarding the nature of prepositions. Some authors consider prepositions to be essentially lexical elements, because they are part of the encyclopedic content of a language

and have different semantics (cf. Marantz 2001, Borer 2003). Baker (2003), on the contrary, considers prepositions functional elements, since they do not have derivational morphology. In this paper, I assume Svenonius' (2004) view that prepositions are basically functional elements, which can also be part of an open category, because new words may be added to the preposition inventory. For instance, it is common to find words that have been grammaticalized into prepositions – for example, *instead of* was grammaticalized from the noun *stead* which means 'place'. Prepositions are thus a hybrid category. To this end, Svenonius proposes the following properties for prepositions (P) based on several aspects in many languages:

- (30) Typical characteristics of adpositions  
 Express binary relations between entities (including events).  
 Form a syntactic constituent with a DP complement.  
 C-select properties of the complement.  
 S-select properties of the complement.  
 Project XPs which function as predicate or sentential adjuncts.  
 Do not combine with tense or aspect morphology.

So as to know if the element is a functional or a lexical preposition, it is necessary to analyze the structure the preposition is part of. For instance, according to (30c), prepositions can c-select elements. C-selection determines syntactic categories, for example, 'during' selects a DP and 'while' a TP or a CP. Additionally, as stated in (30d), prepositions can impose semantic restrictions, hence s-select their complement. For instance, preposition 'on' selects complements with the semantics of *surface*.

According to Svenonius (2004) and Cuervo (2010), as prepositions may have semantic content, they are transitive elements, because they can project complement and specifier in their structure. Additionally, Hale and Keyser (2002) assume prepositions are relational elements that may establish a connection between two arguments.

Given this discussion, I assume that all the prepositions which introduce IAs in BP are transitive prepositions that s-select their arguments and can project a specifier and a complement in the argument structure. In EP, on the other hand, the *a*-DP obligatory alternates with a dative clitic, which is evidence that preposition *a* is a functional element, present in the structure to assign dative Case. In EP, when the preposition has semantic content, it does not alternate with clitics (cf. 22). Given the discussion on prepositions above and the fact that IAs in BP do not alternate with the dative clitic *le* confirm prepositions in BP are lexical transitive elements.

Therefore, the IAs in BP are not introduced by functional prepositions. Additionally, there is no *dative alternation* in BP, as there is no alternation between two structures as exemplified for English and Spanish. Consequently, it is not necessary to postulate applicative heads for BP, as prepositional phrases are perfectly capable to introduce the IAs with transitive prepositions, as it will be discussed in the next section.

## 2.4 The projection pP

In this section, I will demonstrate why a pP projection is better suited for IAs in BP, instead of ApplP and or a simple PP, as it has been proposed for PDCs before. Recall that Svenonius (2004), inspired by Hale and Keyser (2002), establishes that prepositions are relational elements. This relation can be captured through Figure and Ground associations. According to Talmy (1978), the Figure is the moving or conceptually movable object and the Ground is the reference. For example, in ‘John threw the keys on the table’ *the keys* is the Figure, *the table* is the Ground and the element responsible to relate them is the preposition *on*. Therefore, the Ground is the complement of the preposition. Hence, the interpretation of the Ground depends on the preposition, whereas the interpretation of the Figure does not. Thus, as discussed before, transitive prepositions determine selection restrictions on their complement – the Ground – but not on the Figure.

As prepositions can project Spec and complement positions, a pP can be perfectly responsible for introducing a thematic relation between the DO (Figure) and the IO (ground) (cf. Wood 2012). Otherwise, if only a PP projection is accounted for, the DO Theme would sit in SpecPP - being subcategorized by the preposition, when actually its relation is with the verb. Therefore, Wood (2002, 180) proposes a parallel between a pP representation and Voice (cf. 32), insofar as the prepositional structure involves a ‘light preposition’ *p* and P, as categories Voice and *v* in the verbal domain.

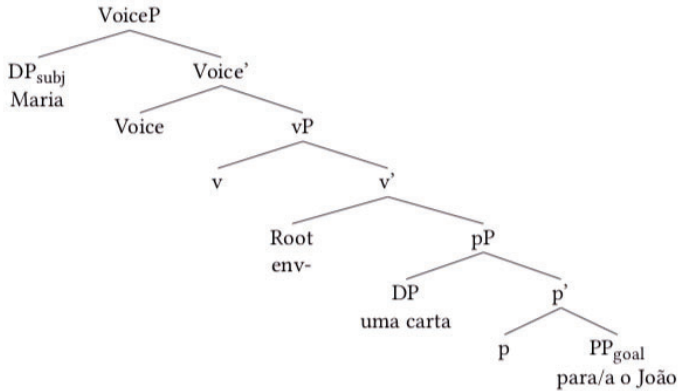
(31) [VoiceP *Agent* [Voice’ [ Voice [vP [v [PP[PTheme[*Goal*]]]]]]]]

(32) [VoiceP *Agent* [Voice’ [ Voice [vP [v [*Theme*]]]]]  
           [pP *Figure* [p’ [p [PP [P [*Ground*]]]]]]

The representations in (32) show the farther relation between the Figure and P, which can be related to the asymmetry concerning the verb and both

its complements in ditransitive constructions. Thus, in a pP configuration, the preposition remains inside the PP, as a preposition imposes restrictions to the Ground (IO), not the Figure (DO). This means *p* can be responsible for introducing a thematic relation. Consequently, dynamic ditransitive structures in BP with *goal* IAs can be represented as follows:

(34)



The representation above solves the questions related to transfer and movement verbs whose IAs are *goals*. In sentences with creation verbs, however, preposition *para* entails two *beneficiary* interpretations, as mentioned before.

In English, when the IA is the *beneficiary of the theme*, it would be represented as a low applicative in Pytkänen's (2002) terms (cf. Section 2.1). For the reasons discussed in Section 2.3, for BP, Armelin (2011) assumes applicative heads for *goals*, PPs for *beneficiaries of the theme* and adjunction for *beneficiaries of the event* (high applicative in Bantu languages).

Returning to the debate on Section 2.1 about Bantu languages, the examples discussed from Chaga and Chichewa were specifically related to *beneficiary* arguments with applicative morphology – *beneficiary of the theme* in Chichewa and *beneficiary of the event* in Chaga. Therefore, in these languages, applicative morphology exists to account for *beneficiary*  $\theta$ -roles. Afterward, the applicative analysis was adapted for English *goals* and *beneficiaries* in double object constructions. Then, ApplP was assumed in some Romance languages. However, when Armelin (2011) proposes applicative heads for BP, this proposal only holds for *goals* but not for *beneficiaries* which were the source for ApplP proposals in the first place.

This discussion leads us to assume the applicative analysis should hold for both *goals* and *beneficiaries*.

However, as discussed in 2.3, BP does not have applicative heads, let alone *high and low applicatives* to capture the different semantics of the *beneficiary of the theme* and the *beneficiary of the event*. Therefore, to represent the two semantic readings preposition *para* entails with creation verbs in BP, it is necessary to find a different approach.

In the next section, I will propose that the i\* single argument introducer provides a solution to represent these structures once we consider the IAs as true arguments that are part of the event. Thus, as they are not adjuncts, an adjunction to VP is not the best representation in this case. As we will see, according to Wood and Marantz (2017), the i\* single argument introducer accounts for the main heads which add participants to the event, namely *Voice*, *low applicative*, *little p*, *prepositions (P)*, and *high applicative*. These options will enable us to capture the different semantics preposition *para* entails.

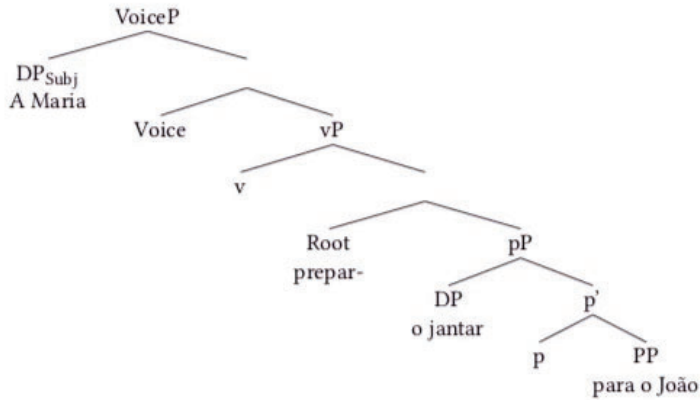
### 3. Representing the two semantic readings for the IA introduced by *para* in BP

Regarding ditransitive sentences with creation verbs, the representation with a pP, similar to the one in (33) is not capable of maintaining the two *beneficiary* interpretations that can be instantiated by *para* with creation verbs in BP. Let's return to example (1), repeated below as (34):

- (34) Maria preparou o jantar para o João / para ele. (BP)  
 Maria prepared DEF dinner P<sub>para(to)</sub> DEF João. OBL/ for him.3P.SG

In sentence (34), the IA can be the *beneficiary of the theme* or *beneficiary of the event*. In the first case, *João* will be the person who will eat the dinner prepared by Maria, therefore *beneficiary of the theme*, which would have the same semantics as the one conveyed by low applicatives in Chichewa (cf. 13) (cf. Pylkkänen 2002), when there is a transfer of possession. In Bantu languages, these structures would have applicative morphology. Since BP does not have applicatives, a possible solution would be a similar representation to (34) with a pP head, as in (35):

(35)



This representation in (35), however, does not capture the two possible semantic readings mentioned before. The second interpretation has the semantics of *high applicatives*, meaning that *João* is not necessarily the person who will eat dinner, he may be the person who had the obligation of making dinner, but could not, hence Maria did it – she is the one who acted in the event – thus, *João* will be the *beneficiary* of Maria preparing dinner in his place, even though he may not be the person who will properly eat it. Hence in the next section, we will discuss the *i\** single argument introducer which gives a solution for representing both semantic readings.

### 3.1 The *i\** single argument introducer

Wood and Marantz (2017) propose a single argument introducer *i\** to account for the main heads which add participants to the event. The authors claim *Voice*, *low applicative*, *little p*, *prepositions (P)* and *high applicative* can be reduced to *i\**. In this section, I will show that BP constructions with *beneficiary* IAs introduced by *para* are very well suited to illustrate how the *i\** can provide an effective way to represent the different semantic readings conveyed by prepositional phrases introduced by *para*. Representation (36) accounts for the IA with low applicative reading – *beneficiary of the theme*, but not for the interpretation of *beneficiary of the event*, which would be instantiated by a high applicative.

According to Wood and Marantz (2017, 260), the high applicative is similar to a root-adjoined it, because the  $\theta$ -role assigned to the DP in SpecAppIP is not conveyed by the vP semantics. Especially with *beneficiaries*,

insofar as the semantics of a vP, as ‘prepare’, do not imply the presence of a *beneficiary*. Therefore, the authors postulate that the  $\theta$ -roles associated with high applicatives are similar to the ones introduced by prepositions, such as *beneficiaries* (English *for*) and *locatives* (English *in*). This assumption is particularly relevant for IAs with creation verbs in BP, which are introduced by *para*. First of all, because they have the semantics of *beneficiary*, second of all, because diachronically, as discussed in Section 1, preposition *para* - that substituted preposition *a* - was primarily used with locatives (cf. 23).

Assuming BP does not have applicative heads in its argument structure, the *beneficiary* IA may be introduced by a prepositional phrase. Firstly because, according to Acedo-Matéllan (2010), prepositions are like any other lexical categories that may have a neutral root, as well as a category that determines the functional heads.

Additionally, prepositions can be prepositional roots with categorial features that will adjoin to  $i^*$  and generate a pP, as we will see next. On the contrary, if the preposition does not have a categorial feature, they assign a  $\theta$ -role to the DP in SpecvP, as a high applicative does. Subsequently, the interpretation of the IA *beneficiary of the event* can be conveyed.

The introducer  $i^*$  is a categorially unspecified head without a categorial feature in the beginning of the derivation. Then, its categorial feature is valued by the categorial feature of the first constituent it merges with from the combination of an unvalued category (CAT), which may or may not trigger Merge, with a constituent of category D, such as: {[CAT:    ], [S: D]} (cf. 36 and 37). The underscore indicates an unvalued CAT feature and  $i^*$  would be the notation for this feature bundle. The selectional features are annotated in brackets, for example, P[S: D]. Therefore, P[S: D] is a head of category P that selects (S) for a constituent of category D (Wood and Marantz 2017: 257).

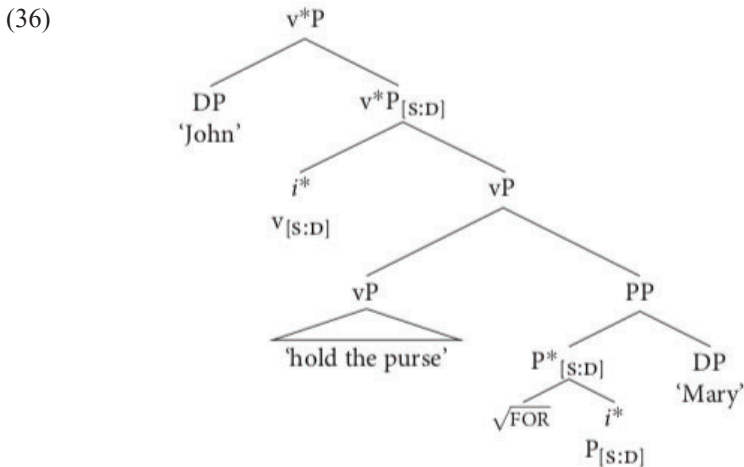
Therefore, the single argument introducer  $i^*$  is valued by the categorial feature of the first constituent it merges with by combining it to an unvalued category, which may or may not trigger Merge (for more details cf. Wood and Marantz 2017, 257). Take the sentence ‘*John held the purse for Mary*’, for instance. The idea of ‘carrying a bag’ does not imply a *beneficiary*. Hence ‘for Mary’ is an extra argument that can be related to  $\theta$ -roles associated with high applicatives or prepositions, as mentioned before.

According to Wood and Marantz (2017), the IA ‘for Mary’ can have two interpretations: a beneficiary PP, when the preposition relates the DP ‘the purse’ to the *beneficiary* ‘Mary’, similar to what I am calling *beneficiary of the theme* someone is carrying the bag in Mary’s benefit. The difference



from a ditransitive sentence, however, is that ‘purse’ is not the Figure of the preposition it merges with - as the OD Theme in a ditransitive would be – but that it is merged in the structure in SpecvP of the verb ‘hold’ (cf. 36).

Thus, the preposition *for* in this case has a categorial feature, which is valued as P because the semantic interpretation of the preposition depends on the prepositional root  $\sqrt{\text{FOR}}$ , hence ‘Mary’ will have the semantics of *beneficiary* in this case. When the lower  $i^*$  merges with  $\sqrt{\text{FOR}}$ , it assigns the DP ‘Mary’ the  $\theta$ -role *beneficiary* associated with it. Therefore, the categorial preposition merges to  $i^*$  and then adjoins the DP ‘Mary’ projecting PP. Next,  $i^*$  merges with vP, valuing its categorial feature as v, and projects  $v^*P_{[S:D]}$  – explicar o S:D, a verb phrase which c-selects an external argument, which is the DP ‘John’, that closes off the complete  $v^*P$ :

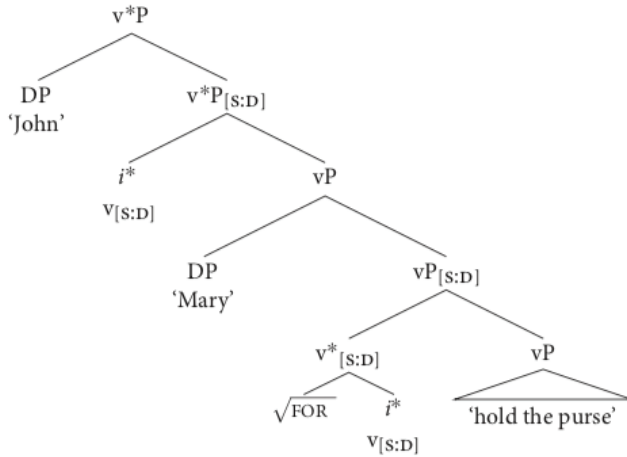


(Wood and Marantz 2017, 261)

Besides the interpretation of *low applicative*, represented in (36), this structure may also have the semantics of a *high applicative*, i.e., *beneficiary of the event*. In this case, there is an event which is ‘carrying the bag’ that is Mary’s responsibility, she is supposed to carry the bag (it may not even be her bag), as she cannot carry it, ‘John’ will do it instead of her, similarly to example (14) from Chaga, discussed previously.

Hence, the preposition instead of merging with the DP and then to the vP, first merges to the vP and then the DP, as a high applicative would do (cf.37):

(37)



(Wood and Marantz 2017, 262)

This representation is possible because, according to Wood and Marantz (2017), in this configuration, a prepositional root may be a neutral category. Therefore, in (38),  $i^*$  first merges with  $\sqrt{\text{FOR}}$ , which has no categorial feature, differently from (36), hence it projects  $v^*$  not  $P^*$ . Next, it merges with  $vP$ , to value its categorial feature  $v$ , subsequently projecting  $vP_{[S:D]}$ . Next, the categorial feature  $D$  is checked when  $DP$  'Mary' merges with  $vP_{[S:D]}$ . Finally, the external  $DP$  'John' is added in the structure.

In the next section, I propose that the theory put forth by Wood and Marantz (2017) may be better illustrated with structures with creation verbs in  $BP$ , insofar as the  $IA$  introduced by *para* can clearly have two interpretations, as mentioned before.

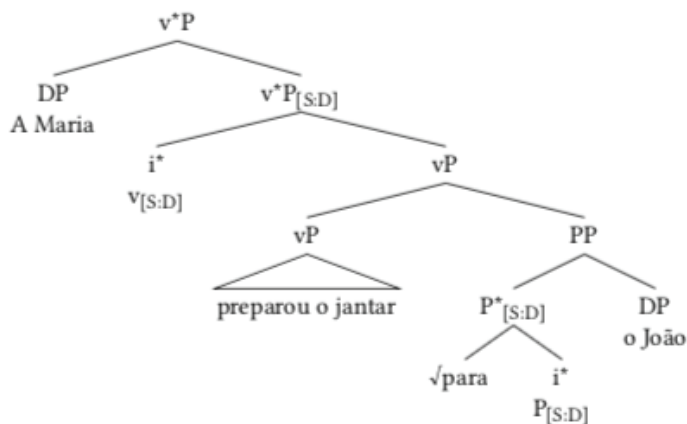
### 3.2 Brazilian Portuguese beneficiaries and the $i^*$ proposal

Applying what was discussed in the previous section to  $BP$ , (39) is the same as (36), and insofar as the preposition *para* has categorial features and the semantic reading of the  $IA$  is *beneficiary of the theme*. Hence, the categorial  $\sqrt{\text{PARA}}$  merges with  $i^*$  and then adjoins to the  $DP$  'o João' projecting a  $PP$ . Subsequently, this constituent adjoins to the  $vP$  'preparou o jantar'. Next, to introduce the external argument 'A Maria',  $i^*$  merges

with vP, valuing its categorial feature as v, projecting a v\* P<sub>[S,D]</sub>, which c-selects a DP ('A Maria'), producing a complete v\*P.

Consequently, in this configuration 'o João' is interpreted as being the *beneficiary of the theme*, i.e., he is the one who dinner was prepared for, as we can assume from the representation below where 'o João' is lower than the vP – recall the discussion presented on *low applicatives* previously.

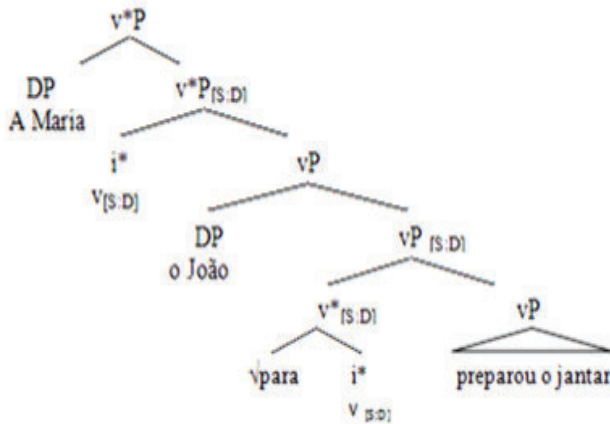
(38)



The example I am focusing on here may have a second interpretation (cf. 39) – dinner may be appreciated by people other than 'o João', which is why 'o João' is the *beneficiary of the event*, i.e., 'o João' is the beneficiary of the event of 'Maria' preparing dinner, he may not even eat it. For instance, perhaps he should prepare dinner, but he is sick, so 'Maria' will prepare it instead of him.

Therefore, in (39), √PARA is a neutral category, hence when merged with i\*, it generates v\*, not P\*. This constituent, when merged with vP 'preparou o jantar' (prepared dinner), values the categorial feature of v by projecting vP<sub>[S:D]</sub>. Subsequently, the categorial feature of D is checked by merging vP<sub>[S:D]</sub> with the DP 'o João' – as in a *high applicative*, where the DP *beneficiary of the event* is higher than the vP. Finally, the external argument 'a Maria' is added to the structure in a similar fashion to (37):

(39)



Therefore, the representations in (38) and (39) capture the different semantics conveyed by *para* in these structures, i.e., *beneficiary of the theme* and *beneficiary of the event*, respectively. As BP does not have applicative heads, Marantz and Wood's (2017) proposal fits perfectly to solve the ambiguous interpretation conveyed by preposition *para* when introducing IAs with creation verbs in BP.

#### 4. Final Remarks

In this paper, I discussed the representation of creation verbs in BP. In the introduction, I showed BP has undergone a diachronic change in terms of the introduction of the IA present in these sentences, from sharing the same structure to the one modern EP still has. In EP, the IA is introduced by preposition *a*, in BP the IA is introduced with *para*.

Following English and other Romance languages, it has been argued EP also has *dative alternation*, as it presents a structure with a functional preposition that alternates with dative clitics (similar to a DOC), and can be introduced in the argument structure by applicative heads.

Additionally, an applicative approach has been proposed to BP. This assumption, however, does not hold, insofar as prepositions in BP are lexical / transitive elements. Therefore, the relation between the DO and the IO selected by the verbal root can be introduced in the argument structure by a projection  $pP$ .

This representation, however, does not capture the two possible semantic interpretations conveyed by the IA introduced by *para* in creation verb structures in BP. Thus, the representation of creation verbs in BP should necessarily involve the single argument introducer *i\**, with which it is possible to provide a more accurate account for both interpretations conveyed by the preposition *para* in these contexts.

## References

- Acedo-Matellán, Víctor. 2010. *Argument Structure and the Syntax-Morphology Interface: A Case Study in Latin and Other Languages*. PhD diss., Universitat de Barcelona.
- Alsina, Alex, and Mchombo, Sam. 1993. "Objects asymmetries and the Chichewa Applicative Construction". In *Theoretical aspects of Bantu grammar*, edited by Sam Mchombo, 50-93. Stanford: CSLI Publications.
- Amaral, Luana, and Cançado, Márcia. 2014. "Verbos de criação do português brasileiro: classificação e representação lexical". *Revista Linguística*. 10 (1): 51-73.
- Armelin, Paula. 2011. "Sentenças Bitransitivas do Português Brasileiro revisitadas à luz da teoria de núcleos funcionais aplicativos". Master's Diss., University of São Paulo.
- Baker, Mark. 1988. *Incorporation: a Theory of Grammatical Function Changing*. Chicago: Chicago University Press.
- Bresnan, Joan, and Moshi, Lioba. 1993. "Object Asymmetries in Comparative Bantu Syntax". *Linguistic Inquiry* 21 (2): 147-185.
- Brito, Ana Maria. 2010. "Do European Portuguese and Spanish have the Double Object Construction?". In *V Encuentro de Gramática Generativa Search for Conference Proceedings Publications*, edited by Facultad de Lenguas-Universidad Nacional del Comahue, 81-114. General Roca, Río Negro: Universidad Nacional del Comahue.
- Borer, Hagit. 2003. "Exo-skeletal vs. Endo-skeletal Explanations: Syntactic Projections and the Lexicon". In *The Nature of Explanation in Linguistic Theory*, edited by Maria Polinsky and John Moore, 31-67. Chicago: Chicago University Press.
- Calindro, Ana. 2015a. "Introduzindo Argumentos: uma proposta para as sentenças ditransitivas do português brasileiro." [Introducing arguments: the case of ditransitives in Brazilian Portuguese]. PhD diss., University of São Paulo.
- Calindro, Ana. 2015b. "Um estudo sobre as preposições introdutoras de argumentos em português brasileiro." [Considerations on the

- prepositions that introduce indirect arguments in Brazilian Portuguese]. *Revista Linguística – Associação de Linguística e Filologia da America Latina (ALFAL)* 31 (2): 61-72.
- Calindro, Ana. 2016. “Introducing indirect arguments: the locus of a diachronic change.” *Rivista di Grammatica Generativa* 38: 35-44.
- Calindro, Ana. 2020. “Ditransitive constructions: what sets Brazilian Portuguese apart from other Romance languages?”. In *Dative structures in Romance and beyond*, edited by Anna Pineda, Jaume Mateu and Ricardo Etxepare, 73-92. Berlin: Language Science Press.
- Carvalho, Janayna, and Calindro, Ana. 2018. “A unified account for the loss of third person clitics in Brazilian Portuguese.” In *Pronomes: morfossintaxe e semântica*, edited by Danniell Carvalho, and Dorothy Brito, 91-110. Salvador: Ed. UFBA.
- Cépeda, Paola, and Cyrino, Sonia. 2020. “Putting objects in order: Asymmetrical relations in Spanish and Portuguese ditransitives”. In *Dative structures in Romance and beyond*, edited by Anna Pineda, Jaume Mateu and Ricardo Etxepare, 97-116. Berlin: Language Science Press.
- Cornilescu, Alexandra, Dinu, Anca, and Tigău, Alina. 2017. “Experimental Data on Romanian Double Object Constructions”. *Revue Roumaine de Linguistique* LXII (2): 157-177.
- Cuervo, Cristina. 2003. “Datives at large.” PhD. diss., MIT.
- Cuervo, Cristina. 2010. “Against Ditransitivity”. *Probus* 22 (2): 151-180
- Diaconescu, Constanta Rodica, and Rivero, María Luisa. 2007. “An applicative analysis of double object constructions in Romanian”. *Probus* 19 (2): 209–233.
- Figueiredo Silva, Maria Cristina. 2007. “A perda do marcador dativo e algumas das suas consequências.” [The loss of the dative marker and some consequences]. In *Descrição, história e aquisição do português brasileiro*, edited by Ataliba de Castilho, Maria Aparecida Torres Morais, Ruth Lopes, and Sonia Cyrino, 85-110. Campinas: Ed. Pontes.
- FLiP. n.d. “Gramática – preposição.” [Grammar - preposition]. Accessed January 20, 2020.  
<https://www.flip.pt/FLiP-On-line/Gramatica/Morfologia-Partes-do-discurso/Preposicao>.
- Galves, Charlotte. 2001 *Ensaio sobre as gramáticas do português*. Campinas: Editora da Unicamp.
- Guéron, Jacqueline. 1985. “Inalienable possession, PRO-inclusion and lexical chains.” In *Grammatical Representation*, edited by Jacqueline

- Guéron, Hans-Georg Obenauer, and Jean-Yves Pollock, 43 – 86. Dordrecht: Foris.
- Hale, Ken and Keyser, Samuel. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, Mass.: The MIT Press.
- Kato, Mary, Sonia Cyrino, and Vilma Corrêa. 2009. “Brazilian Portuguese and the recovery of lost clitics through schooling.” In *Minimalist Inquiries into Child and Adult Language Acquisition*, edited by Acrisio Pires and Jason Rothman, 245-272. Mouton de Gruyter.
- Larson, Richard. 1988. “On the double object construction”. *Linguistic Inquiry* 19 (3): 335-391.
- Levin, Beth. 1993. *English verb classes and alternations: a preliminary investigation*. Chicago: University of Chicago Press.
- Levinson, Lisa. 2007. “The roots of verbs.” PhD diss., New York University.
- Marantz, Alec. 1993. “Implications of asymmetries in double objects constructions”. In *Theoretical aspects of Bantu grammar*, edited by Sam Mchombo, 113-151. Stanford: CSLI Publications.
- Marantz, Alec. 2001 “Words”. Paper presented at *20th West Coast Conference on Formal Linguistics*, University of Southern California, February 23-25.
- Marantz, Alec. 2007. “Phases and Words”. In *Phases in the Theory of Grammar*, edited by Sook- Hee Choe, Dong-Wee Yang, Yang-Soon Kim, Sung-Hun Kim, and Alec Marantz, 191–222. Seoul: Dong-In Publishing Co.
- Mazurkewich, Irene. 1984. “The acquisition of the dative alternation by second language learners and linguistic theory”. *Language Learning* 34: 91-110.
- Pineda, Anna. 2013. “Double object constructions in Spanish (and Catalan) revisited”. In *Romance Languages and Linguistic Theory 2011: Selected Papers from ‘Going Romance’*, edited by Sergio Baauw, Frank Drijkoningen, Luisa Meroni, and Manuela Pinto, 193-216. Amsterdam: John Benjamins.
- Pinker, Steven. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.
- Primeira página: 90 anos de história nas capas mais importantes da Folha*. 2011. [90 years of History through the most important covers of Folha]. Folha de São Paulo. 7a ed. São Paulo: Publifolha.
- Pujalte, Mercedes. 2010. “Construcciones com y cin aplicativos em el español del Rio da la Plata y el português brasileiro”. Ms., Universidad Nacional del Comahue.

- Pylkänninen, Liina. 2002. *Introducing Arguments*. PhD. diss., MIT.
- Scher, Ana Paula. 1996. *As Construções com Dois Complementos no Inglês e no Português do Brasil*. [Constructions with two complements in English and Brazilian Portuguese]. Master's diss., Unicamp.
- Svenonius, Peter. 2003. "Limits on P: filling in holes vs. falling in holes". *Nordlyd* 31:431-445.
- Svenonius, Peter. 2004. "Adpositions, particles and the arguments they introduce". In *Argument Structure*, edited by Eric Reuland, Tanmoy Bhattacharya, and Giorgos Spathas, 63-103. Philadelphia: John Benjamins.
- Talmy, Leonard. 1978. "Figure and ground in complex sentences". In *Universals in human language*, edited by Joseph Greenberg, 625-649. Stanford University Press.
- Torres Morais, Maria. 2007. *Dativos*. [Datives]. Professorship Thesis, University of São Paulo.
- Torres Morais, Maria, and Berlinck, Rosane. 2007. "'Eu disse pra ele' ou 'Disse-lhe a ele': A expressão do dativo nas variedades brasileira e europeia do português". ['Eu disse para ele' or 'Disse-lhe a ele': the dative expression in Brazilian and European Portuguese varieties]. In *Descrição, história e aquisição do português brasileiro*, edited by Ataliba Castilho, Maria Aparecida Torres Morais, Ruth Lopes and Sônia Cyrino, 61-74. Campinas: Ed.Pontes.
- Torres Morais, Maria, and Berlinck, Rosane. 2018. "O objeto indireto: argumentos aplicados e preposicionado". ["The indirect object: applied and prepositional arguments"]. In *Mudança sintática do português brasileiro: perspectiva gerativista*, edited by Sônia Cyrino, and Maria Aparecida, 252-307. São Paulo: Contexto.
- Torres Morais, Maria, and Salles, Heloísa. 2010. "Parametric change in the grammatical encoding of indirect objects in Brazilian Portuguese." *Probus* 22 (2): 181-209.
- Wood, Jim. 2012. "Icelandic Morphosyntax and Argument Structure." PhD diss., New York University.
- Wood, Jim, and Marantz, Alec. 2017. "The Interpretation of External Arguments." In *The Verbal Domain*, edited by Roberta D'Alessandro, Irene Franco and Ángel Gallego, 255-278. Oxford University Press.



## Notes

<sup>1</sup> I would like to thank an anonymous reviewer for detailed and helpful comments, and suggestions to an earlier version of this paper.

<sup>2</sup> According to Brito (2009) *para* is also a possibility to introduce the *beneficiary* with creation verbs in European Portuguese. For now, my goal is the structure in BP, leaving these questions related to EP for future work.

<sup>3</sup> The context of 1<sup>st</sup> and 2<sup>nd</sup> person is different from 3<sup>rd</sup> person for a number of reasons as discussed in Carvalho and Calindro (2018). In this paper, I will discuss the facts related to the 3<sup>rd</sup> person, following the works on other languages, which focus mainly on 3<sup>rd</sup> person as well, as Cuervo (2003), for Spanish.

<sup>4</sup> Levinson (2007) accepts both implicit and explicit creation verbs. In this paper, I consider only the explicit group, because this duality *implicit / explicit creation verbs* does not hold in BP.

<sup>5</sup> In order to have a coherent quantity of data among decades, the number of words in each decade was taken into account. Hence, the number of words analyzed from each decade varies between 22,000 to 28,000, in a total of 235,587.

<sup>6</sup> Amaral and Caçado (2014, 52) attest there is a lack of quantitative studies regarding creation verbs with diachronic data in BP. Hence, in the future, it will be important to continue the research based on different corpora more suited to find creation verbs.

<sup>7</sup> The dates in parentheses refer to the *Folha de São Paulo* front page the example was taken from.

<sup>8</sup> I would like to thank an anonymous reviewer who called my attention to the importance of presenting the exact number of sentences per century.

<sup>9</sup> To these verbs with Latin roots, Larson (1988) adds *donate* and *distribute* (John donated money **to charity**; \*John donated **charity** the money). There is, however, another asymmetry, when the DOC does not alternate with the PDC, as with *spare* and *envy* (The judge spared **John** the ordeal \*The judge spared the ordeal **to John**).

<sup>10</sup> Abbreviations used by Marantz (1993, 115) SP = Subject prefix (subject agreement); pst = past INFL; APPL = applicative affix (afixal verb) ; fv = final vowel; prs = present INFL.

<sup>11</sup> In the high applicative tree (15), Pylkkänen (2002, 19) uses English words instead of Chaga. As I am using this tree to illustrate how a high applicative is represented, I chose to keep the author's syntactic tree in English.

<sup>12</sup> I would like to thank an anonymous reviewer who pointed out the data presented in Diaconescu and Rivero (2007) are not entirely accurate. Cornilescu, Dinu and Tigău (2017) conducted an extensive and thorough experiment, which shows clitic doubling of the IA is not obligatory in all Romanian varieties. Therefore, as the presence of the pronominal clitic is optional, the authors assume Romanian does not have DOCs. I intend to return to this very interesting study in the future, as the discussion presented relates to the study conducted by Cépeda and Cyrino (2020) on Spanish, EP and BP, who do not assume DOCs as well. Additionally, Pineda

(2013) also argues Spanish has DOCs, even though, she does not assume clitic doubling is obligatory. Hence, in this paper, I will maintain Diaconescu and River's approach, because my main aim is to show that the clitic plays an important role for authors to assume or not *dative alternation* in Romance languages. Therefore, the fact that BP has lost 3rd person clitic *lhe(s)* is a relevant evidence for not assuming applicative heads in its argument structure.

<sup>13</sup> According to Diaconescu and Rivero (2007), there are, actually, four syntactic versions of ditransitive constructions in Romanian. The ones presented in (19) and (20) and also, there is what is called by the authors a *bare Pdative construction* – *Mihaela trimite la Maria o scrisoare* (Mihaela sends to Mary a letter), whose interpretation is similar to (20). Additionally, there is a *clitic doubled Pdative construction*, which has been documented, in the literary works of Transylvanian authors, but it is not part of Romanian prescriptive grammar – *Mihaela îi trimite la Maria o scrisoare* (Mihaela sends Mary a letter). These are very interesting data, however, for the purpose of this paper, examples (19) and (20) suffice.

<sup>14</sup> *Zona da Mata, Minas Gerais* state on the border with the states of *Rio de Janeiro* and *Espírito Santo*.

<sup>15</sup> According to Marantz (1997), verbs are formed in the syntax by the combination of a lexical root and a verbalizing head little *v*.

<sup>16</sup> For a comprehensive study on dative complements in BP cf. Figueiredo Silva (2007).

## CHAPTER FOUR

### DATIVES IN ISTRO-ROMANIAN

IONUȚ GEANĂ

This paper deals with a particular phenomenon of Istro-Romanian morphosyntax, namely the realization of the Dative. After describing the place of Istro-Romanian (IR) across (Eastern) Romance, this article describes the way IR marks its cases: nominative-accusative vs dative-genitive syncretism for nouns, (suppletive) special forms for each case for pronouns. After a careful analysis of IR corpus, I will answer three questions: (i) is the IR dative morphology and use (any) different from standard Romanian?; (ii) Does IR dative pattern with any other (Eastern) Romance language or variety?; (iii) Is IR dative system innovative in any way? The conclusions will show that each question brings in complex answers for this understudied Romance variety.

Keywords: Eastern Romance, Istro-Romanian case marking, synthetic datives, analytical datives, the Benefactive

#### 1. Introduction

In this paper, I will give an account of the functioning of the dative case in Istro-Romanian (IR). Before doing so, I will describe in this introductory section the place of Istro-Romanian across (Eastern) Romance, showing some features of Istro-Romanian mainly in comparison with Daco-Romanian. This section also sets my three research questions with focus on IR datives:

1) Is the IR dative morphology and use (any) different from standard Romanian?

2) Does IR dative pattern with any other (Eastern) Romance language or variety?

3) Is IR dative system innovative in any way?

Other than the Introduction and Conclusions, my paper has three main sections, focusing on: I. IR Case marking, II. IR Datives, III. Double-

Marked IR Datives. At this point, although my three research questions apparently require a straightforward yes or no answer, I will show that the answer is more complex and nuanced than that, each of the questions leading to at least twofold answers.

The study of Istro-Romanian as a Romance variety (hereinafter referred to as IR) has not been the focus of thorough research until the 19<sup>th</sup> century, through research carried out by Maiorescu 1874, consistently continuing in the 20<sup>th</sup> and 21<sup>st</sup> c. (for recent work and research on IR, see Geană 2017a, 2017b, 2018). Istro-Romanian is currently spoken in two (once large) groups: the northern group – centred around the village of Žejane, and the southern group – made of several villages, most prominently Šušnjeвица. These two areas are separate by around 50 km (an administrative border today). Istro-Romanians live in today's Croatia (in an area that used to belong to Italy), in a multi-ethnic and multilingual environment; language contact – especially with Croatian – is widely attested and accepted (for details, see Caragiu Marioțeanu (1977: 213-5), Kovačec (1984: 550-4), Vrzić & Singler (2016: 51)).

Istro-Romanian is an Eastern Romance variety, currently severely endangered, according to the UNESCO *Atlas of the World's Languages in Danger*, and it is spoken today mostly in Croatia by around 500 people, being thus the least spoken Eastern Romance variety<sup>1</sup>. There are no public institutions that use this variety, and, apart from some proverbs and song fragments, there is barely any Istro-Romanian folkloric literature. Their national awareness is rather vague. However, starting from their linguistic particularities, they are glad to show they are different from Croatians, and may call themselves *rumuni*, or, in the south, *vlaš*. Despite the fact that there are two areas where Istro-Romanian is spoken, those from the north have little to no contact with those from the south, and mostly have no feeling of sharing the same origin. The constant bilingualism (Istro-Romanian and Croatian) led to Croatian having a strong influence on Istro-Romanian. Apart from Croatian, according to the region where Istro-Romanian villages are (or were) present, claims have been made on the influence of Italian and/or Slovenian.

In line with the official acceptance of the Romanian Academy (Rusu 1984), Romanian includes four historical dialects: Daco-Romanian, Aromanian, Megleno-Romanian, and Istro-Romanian. The north-Danubian dialect (from Romania and the Republic of Moldova) is dialectologically known as Daco-Romanian (which I will refer to as standard Romanian), spoken as a mother tongue by more than 20 million people (figures vary according to source, ranging from 20 to 29 million). The sub-Danubian dialects (Aromanian, Megleno-Romanian, and Istro-Romanian) form

compact zones in Greece, Albania, Northern Macedonia, and Croatia. There are approximately 700,000 Aromanian speakers, 6,500 Megleno-Romanian speakers and a little over 100 Istro-Romanian speakers (Vrzić 2019, *mss*). In particular regard to Istro-Romanians, their belonging to an ethnic group or their speaking the same variety vary to a great extent, hence the differences in census numbers. The status as a language or a dialect is usually a problem among Romanian linguists and dialectologists (see, for example, Coteanu 1957).

There are some features that individualize the Istro-Romanian in comparison with Daco-Romanian (Kovaceč 1984), briefly mentioned here:

- the rhotacism of intervocalic *n* (as in the north-western part of the Daco-Romanian territory): *mâre* ‘tomorrow’, *bire* ‘well’, *pâre* ‘bread’;
- the labials before palatal vowels are not palatalized: *bire* ‘well’, *pičór* ‘leg’;
- after labials, *e* and *ea* (>*ɛ, ǎ*) do not change to *ǎ*, respectively *a*: *per* ‘hair’, *pɛna* ‘feather’, *fe̯ta* ‘girl’;
- the *i* vowel does not change before a nasal or a labial: *cuvintâ* ‘to speak’, *vint* ‘wind’;
- the monophthongs *ɛ* and *o* correspond to diphthongs *ɛa* and *ɔa*: *sɛra* ‘evening’, *nópte* ‘night’;
- the preservation of the constant clusters *cl*, *gl*: *cl'emâ* ‘call’, *ɣl'âța* ‘ice’, and also the consonants *l* and *n*: *fil* ‘son’, *spuñ* ‘I say’.

## 2. The IR Case System

The Istro-Romanian case system follows the Eastern Romance (and other Balkan languages) of opposing Nominative-Accusative to Dative-Genitive (but see the literature below). The Vocative has forms of its own (like in standard Romanian), nevertheless the Nominative can equally be used with the same function as the vocative.

Just like in other (Eastern) Romance, the case system of pronouns is richer than the noun system, personal pronouns showing a full case paradigm.

Generally, all IR literature mentions the genitive-dative syncretism in IR:

- Kovačec 1984 (genitive and dative are treated together at the subchapter for *Articles*, under declension)

As far as nouns are concerned, in the south, *lu* is used before the NP/DP – “article-preposition”, genitive-dative marker for all genders and numbers,

and feminine singular nouns (and adjectives) in ending in *-e* (in the nominative-accusative) can alternate with apparently no difference in meaning or context with *-e* (the equivalent of a somehow marked genitive in standard Romanian):

- (1) *lu bābe* (DAT *lu bābe* (DAT *lu bāba* (DAT  
old.lady.DAT) old.lady) old.lady.DEF)  
*lu vāke* (DAT *lu vāke* (DAT cow) *lu vāca* (DAT cow.DEF)  
cow.DAT)  
'to/for the old lady'<sup>2</sup> / *babei*  
'to/for the cow' / *vacii*

In the north, all feminine nouns switch to *-e* in the dative if no definite article in postposition is used; in the singular: *lu* for masculine nouns, *le* for feminine nouns, and rarer *-lui*, *-lei*; in the plural: *-lor* for all genders. If the noun is preceded by an adjective or an adjective is nominalized, then *a* is used to mark the dative for all persons and genders (Kovaceč 1984: 568), see below all examples from Žejane:

- singular: masculine in (2a), feminine in (2b):

- (2) a. *Ām* *zis* **omului** SF  
46  
have.1SG.AUX tell.PPLE man.DEF-DAT  
'I told the man' / *Ām zis omului*  
b. *Hobo... zis-a* **mul'ār'ei** a *lu*  
*tātu* TC 120  
*Hobo* tell.PPLE=has.AUX wife.DEF-DAT GEN  
*thief.DEF*  
'Hobo told the thief's wife' / *Hobo i-a zis muierei hoțului*

-plural: feminine in (3a), masculine in (3b), and coordinated nouns in (3c):

- (3) a. *Cuvintāt-a* **măcichelor.** TC 130  
speak.PPLE=has.AUX cats.F.DEF-DAT  
'He told the cats' / *A zis mățelor*  
b. *Hobo... av zis* **tăților.** TC 119  
*Hobo* has.AUX tell.PPLE thieves.DEF-DAT  
'Hobo told the thieves' / *Hobo a zis hoților*  
c. *Av zis* **cohilor** și **camaierilor**  
TC 122  
has.AUX tell.PPLE cooks.DEF-DAT and valets.DEF-DAT

‘He told the cooks and the valets’ / *A zis bucătarilor și cameriştilor*

As far as pronouns are concerned, they are stressed and non-stressed clitics, similar to standard Romanian (Kovačec 1984: 568) – but, I would add, with a different distribution, to which we will get back shortly:

- *ti* in the south instead of *Ț* in the other areas; *t*’ all across IR (Kovačec 1984: 569);

		1 <sup>st</sup> person	2 <sup>nd</sup> person	3 <sup>rd</sup> person		
				Reflexive	Non-reflexive	
					masculine	feminine
Singular	Nominative	}ó	tú	/	}é	}â
	Dative	mî}e	tî}e	sî}e	a lú}	a l’é}
	Accusative	míre	tíre	síre	}é	}â
Plural	Nominative	nó}	vó}	/	}él’	}âle
	Dative	a nó	a vó	sî}e	a lór	
	Accusative	nó}	vó}	síre	}él	}âle

**Table 1. Non-stressed forms (pronominal clitics)** (Kovaceč 1984: 567)

		1 <sup>st</sup> person	2 <sup>nd</sup> person	3 <sup>rd</sup> person		
				Reflexive	Non-reflexive	
					masculine	feminine
Singular	Dative	âm	âț	âș	âp’	
	Accusative	me	te	se	âp’	o (vo)
Plural	Dative	na	va	âș	la	
	Accusative			se	âp’	le

**Table 2. Stressed forms** (Kovaceč 1984: 572)

According to Caragiú-Marioțeanu, Giosu, Ionescu-Ruxăndoiu, Todoran (1977), (hereinafter called *Manual*), IR noun declension includes the following:

- case oppositions have almost completely disappeared (*Manual* 219), with the “apparent” exception of *Žejane*, opposing *-a* in N/ACC to *-e* in D/G;
- *lu* is the definite article for the genitive and dative of all nouns for both numbers and genders;
- however, when describing the definite articles, they are told to be preponderantly proclitic in the D/G singular and plural;
- *Manual* 220: there are a series of articles that vary in gender and preceded by an invariable *a* (more frequent in northern IR), as in Aromanian (AR):

- (4) a. Av zis **a lu** **tatu** (*Manual* 220) (IR)  
 they.have.AUX say.PPLe DAT thief.DEF  
 ‘They told the thief’ / *Au zis hoțului*
- b. Fiŭ pișe a lu ceăia TC 114  
 son.DEF wrote DAT father  
 ‘The son told his father’ / *Fiul a spus tatei*
- c. Atunce av ceŭi doi fraț zis  
**a lu Hobo** TC 117  
 then have.AUX DEF two brothers tell.PPLe  
 DAT Hobo  
 ‘Then the two brothers told Hobo’ / *Apoi cei doi frați i-au spus lui Hobo*
- d. Zmăiu **a le** **măie** spus-a TC 129  
 dragon.DEF DAT mother tell.PPLe=has.AUX  
 ‘The dragon told the mother’ / *Zmeul i-a spus mamei*
- (4‘) L’-đisirî **al<sup>u</sup>** **fur<sup>u</sup>**  
*Manual* 220 (AR)  
 CL.M.DAT.3SG=said DAT thief.DEF  
 ‘They told the thief’ / *Îi ziseră hoțului*
- (5) Av ajutat **a le** **măie**  
*Manual* 220 (IR)  
 they.have.AUX help.PPLe DAT  
 mother.UNMARKED
- (5‘) L’-ađutări **ali** **măii**  
*Manual* 220 (AR)  
 CL.F.DAT.3SG=helped DAT mother.DEF.DAT  
 ‘They helped mother’ / *I-au ajutat mamei = Au ajutat-o pe mama*

- the enclitic article in the D/G is very uncommon;
- no specific notes on the use of personal pronouns in D/G;

SF (1998: 22) claim that the genitive and the dative are both synthetic and analytic, making no note on their frequency in the Romanian version (page 22), but stating in the English version (page 39) that “the analytical forms [...] are widely spread. Synthetic forms can be found only in poetry, sayings and proverbs, where these forms became fixed long ago”;

Zegrean (2012: 31-34) follows the Balkan and Eastern Romance literature.

Here is a simplified table of case of Istro-Romanian nouns with a definite article:



Case	Gender	Singular	Plural
N≡ACC	M	-u, (-u), -le	-i, -le
	F	-a, -vu	-le
G≡D	M	(a) lu(i), (a) -lu(i), a	(a) -lor, a
	F	(a) le, lu, -lei, a	(a) -lor, a

Table 3. The IR definite article

### 3. Istro-Romanian Datives

For the purposes of this article, I will focus on the Dative, comparing it to standard and regional Romanian. Among common uses in standard Romanian and Istro-Romanian, the IR dative is used:

- to express the indirect object (masculine noun in 6a, 6a', feminine noun in 6b, 6b', and pronouns in 7a,b):

(6) a. Hlăpețu lu cesăru ... zice lu gospodăru  
TC 5

guard.DEF GEN emperor.DEF says DAT gentleman.DEF  
'The emperor's guard tells the gentleman' / *Garda împăratului îi spune domnului*

a'. Ganescu lu Mărtin. TC 18

tell.PS.1SG DAT Martin

'I'm telling Martin' / *Îi spun lui Martin*

b. Cea fraierițe ganę lu cea băbe TC 7  
DEF fiancée told DAT DEF old.lady

'The fiancée told the old lady' / *Logodnica îi spuse babei*

b'. Mărtin... ganę lu muļera. TC 19

Martin told DAT woman.DEF

'Martin told the woman' / *Martin i-a zis muierii*

(7) a. Cea ista nu daien ție, ni lu  
nici ur TC 6

DEF this.F.SG NEG give.PS.1PL you.SG.DAT nor DAT  
no one

'We won't give this one [girl] to you, nor to anyone else' / *Pe asta nu ți-o dăm nici ție, nici altcuiva*

b. Zițe prevtu a pei SP 20

say.PS.3SG priest.DEF DAT her.DAT

'The priest tells her' / *Preotul îi zice*

- dative clitic pronoun to double a topicalised indirect object:

- (8) **Lu** **cela** **uâtu,** **ce-a**  
 fost la spine,   
 DAT DEF.M.SG other.DEF who=has  
 been at cork  
**âl'** ganę. TC 11  
 CL.M.DAT.3SG told  
 'He told the other one, who was at the cork' / *Celuilalt, care a fost la dop, i-a spus*

- in possessive dative construction:

- (9) a. **Pure-m** *uocli!* TC 8  
 Put.IMPERATIVE=CL.POS-DAT.1SG eyes.DEF  
 'Put my eyes' / *Pune-mi ochii*  
 b. **Și cum li s-âv muiâra**  
 kemât? SF 48  
 and how CL.POS-DAT.M.3.SG CL.REFL.3.SG=have.AUX wife.DEF  
 called  
 'And what was his wife's name?' / *Și cum i se chema muierea?*

Shifting now towards pronominal clitics, it can be noticed that the 2<sup>nd</sup> person singular dative clitic can vary phonologically to a high degree, as in:

- (10) a. **T-oi** spure nuște TC 6  
 you.DAT.2SG=will.AUX say.INF something  
 'I will tell you something' / *Ți-oi spune ceva*  
 b. **Ț-oi** rată c-er  
 face TC 37  
 you.DAT.2SG=will.AUX show.INF what=COND  
 do.INF  
 'I'll show you what to do' / *Ți-oi arăta ce-oi face*  
 c. **Nu ti-i sete?** TC 11  
 NEG you.DAT.2SG=is thirst  
 'Aren't you thirsty?' / *Nu ți-e sete?*

However, for the southern Istro-Romanian, I have identified the accusative-dative syncretism in the case of 2<sup>nd</sup> person singular pronominal clitic (like Western Romance, but unlike DR), compare (10a) above to (10') below:

- (10') Cum **t-oi** io cea votę conoște?  
 TC 39  
 how you.ACC.2SG=will.AUX I that time know  
 'How will I know you then?' / *Cum te-oi cunoaște atunci?*

Syncretic forms are also attested for 1<sup>st</sup> person Dat=Acc clitic for southern IR (in and around Šušnjeвица). Although it is not specified in the theoretical description, the glossary to SF 225 labels *m* as both a dative and an accusative clitic (not specifically, but recoverable from their glosses):

- (11) Pure-**m** uoc!i! TC 8  
 Put.IMPERATIVE=CL.POS-DAT.1SG eyes.DEF  
 'Put my eyes' / *Pune-mi ochii*
- (11') Preftu **m-a-ntrebât** TC 111  
 priest.DEF CL.ACC.1.SG=has.AUX=ask.PPLE  
 'The priest asked me' / *Preotul m-a întrebat*

The third person singular dative-accusative syncretism is registered in Kovačec (1984: 572). Thus, in southern IR, we can talk – at least for certain contexts – about a full paradigm of dative-accusative syncretic clitics.

Moreover, southern IR displays allomorphic variation, compare (12a) and (12b):

- (12) a. Mărtin **le** ganę. TC 18  
 Martin CL.DAT.3PL told  
 'Martin told them' / *Martin le zise*
- b. Lucifer ganę **lor**. TC 19  
 Lucifer told them.DAT  
 'Lucifer told them' / *Lucifer le zise*

Therefore, can it be assumed that there is an enclisis/proclisis asymmetry, à la Pescarini 2018? The answer is that such asymmetry does not seem to exist, as pronouns seem to be in free variation, see:

- (13) Fraieriță **lui** ganę TC 7  
 fiancée.DEF DAT.3SG told  
 'His fiancée told him' / *Logodnica îi spuse*
- (14) Ie-**P** ganę TC 6  
 he=CL.DAT.3SG told  
 'He told her' / *El îi spuse*

This free variation is noticeable also for 1<sup>st</sup> person singular pronouns:

- (15) a. Nu            **mń-a**            niş            dǎt TC 23  
 NEG            me.DAT=has.AUX nothing give.PPLE  
 ‘He gave me nothing’ / *Nu mi-a dat nimic*
- b. Júta-**mi**,                                  un            centesimo            se  
     uári TC 22  
 help.IMPERATIVE=me.DAT one penny if  
 have.2SG  
 ‘Help me with a penny, if you have’ / *Ajută-mă, un ban dacă ai*  
 BUT
- c. Spure                                  tu            **miie**                                  ce-i cu            tire? TC 25  
 say.IMPERATIVE you me.DAT what=is with you.ACC  
 ‘Tell me what’s wrong with you’ / *Spune-mi ce-i cu tine?*
- d. **Miie**            fost-a                                  sila TC 40  
 me.DAT be.PPLE-has.AUX hurry.NOM  
 ‘I was in a hurry’ / *Eram pe grabă*
- e. Nu            va                                  **miie**            dǎ TC 90  
 NEG will.AUX me.DAT give.INF  
 ‘He will not give [it] to me’ / *Nu-mi va da*

Does clitic doubling occur in IR? Let us take a closer look at the following examples:

- (16) Guárdiia            **P-a**    dǎt            cela            list  
 guard.DEF CL.DAT.M.3SG=has.AUX give.PPLE DEF  
     letter  
**lu cela            camaier**                                  TC 13 (southern IR)  
     DAT DEF valet  
 ‘The guard gave the letter to the valet’ / *Garda i-a dat scrisoarea*  
*cameristului*
- (17) **E-av**    zis                                  **lu            cela**  
     **mladichi** TC 25  
 CL.DAT.M.3SG=has.AUX tell.PPLE DAT DEF  
     young.man  
 ‘He told the young man’ / *I-a zis tânărului*
- (18) Čela            **mi-e**                                  măi            bur            **miie**. SF 74 (north  
 IR)  
 that me.DAT=is more good me.DAT  
 ‘That one is better for me’ / *Acela mi-e mai bun*

- (19) Ie            le                            gaņ        lor.  
 TC 89  
 he        CL.DAT.3PL        told        them  
 ‘He told them’

#### 4. Is there a Double-Marked Dative in IR?

Let us take a look at all these sentences attested for southern IR. My guess is that in all these sentences, due to dative-genitive syncretism, southern IR reorganized its case system, using the preposition *za* (of Croatian origin) to additionally mark a dative noun. As a matter of fact, since the IR nouns in the dative are essentially unmarked morphologically, we can extend the nominative-accusative syncretism to include datives as well. The preposition *za* assigns Accusative, so its combination with the (until now) datives raises a series of problems of interpretation in terms of case. From this perspective, we can say that southern IR patterns with sub-standard Romanian (and the majority of Romance languages and varieties), where analytical datives are pretty common: *la* and even *lu-a* determiner originally—for all genders and numbers.

- (20) Țela            porc        a            facut                            ăń        boske  
       ză-se  
 DEF            pig        has.AUX make.PPLE        in        forest  
       for=himself  
 o        camereș,            **za lui**        **măie**        ățe.        SP 57  
 a        room                    DAT        mother        another  
 ‘The pig made a room in the woods for himself, and another one for his mother’ / *Porcul și-a făcut în pădure o cameră sieși și alta, maică-sii*
- (21) Va            cumpără **za lu**        **șe**            **fil’e** TC 8  
 will.AUX buy.INF    DAT        her        daughter  
 ‘She will buy for her daughter’ / *Îi va cumpăra fiică-sii*
- (22) Ren                            făce        bire        **za lu**        **gospodăru** TC 78  
 would.1PL            do.INF        well        DAT        gentleman.DEF  
 ‘We would make the gentleman a favour’ / *I-am face un bine domnului*

- (23) Oștaru nepișeit-a listu mușât  
**za lu cesăru** TC 13  
 publican.DEF write.PPLE=has.AUX letter.DEF nice  
 DAT emperor.DEF  
 ‘The publican wrote the nice letter for the emperor’ / *Cârciumarul i-a scris o scrisoare împăratului*
- (24) Ie viche che... va **za lu tustrei** fi TC 72  
 she cried that will.AUX DAT all.three be.INF  
 ‘She cried what would happen to the three of them’ / *A plâns despre ce li se va întâmpla celor trei*
- (25) Nu-i bire za me, ni **za lu voi**  
**doi** TC 25  
 not=is good DAT I.DAT nor DAT you.PL  
 two  
 ‘It’s good neither for me, nor for you two’ / *Nu-i bine nici pentru mine, nici pentru voi doi*
- (26) Na mesto de zețe soldi **za lu maiche boije**, a  
 zis TC 102  
 in place of ten coins DAT mother god has.AUX  
 say.PPLE  
 ‘Instead of [giving] ten coins for Virgin Mary, he said’ / *În loc de zece bani pentru Maica Domnului, a zis*
- (27) Uâm ăntrbât pre domnu che se va  
 have.AUX ask.PPLE DOM sir.DEF that if will.aux  
 dă cărstu **za lu țiățe**. TC  
 give.INF christening.DEF DAT father  
 ‘I asked God if he would christen my father’ / *L-am întrebat pe Domnul dacă-i va da botezul tatei*

Note in example (27) that, although the literature claims there is no DOM in IR, there are examples that attest this phenomenon for IR as well (here and in other sources), which seems to be absent from Aromanian and Megleno-Romanian. Other attested IR examples in (28) and (29):



- analytical datives with *lu, a lu, za lu*; but similar to spoken Romanian

Moving on to my second question (*Does IR dative pattern with any other (Eastern) Romance language or variety?*), the answer is again yes and no. Yes, because at least in part dative-accusative opposition is neutralized, and IR patterns with Aromanian for *a lu / al<sup>u</sup>* dative marking; and no, because the patterns are inconsistent (explainable given that IR is not standardized).

And, last but not least, the answer to the third question – *Is IR dative system innovative in any way?* – is yes, considering the southern IR dative realization with *za lu*, a construction not mentioned in previous literature. At least from the data I had access to, the particularity of the *za lu* dative is unique (and novel) across Romance, in general, and Eastern Romance, in particular (and even across IR varieties), given that a preposition (of Croatian origin) combines with a pronominal form (at origin) to mark a dative, whereas in other varieties, for example DR, the dative is formed either synthetically (with an enclitic or proclitic *lui* form), or analytically (substandard) with the preposition *la*, but never combining *la* and *lui*.

## Sources

- SF Sârbu, Richard, and Vasile Frăţilă. 1998. *Dialectul istroromân*. Timişoara: Armacord.
- SP Puşcariu, Sextil. 1906. *Studii istroromâne. Texte*. Bucureşti, Cultura Naţională, Institutul de Arte Grafice „Carol Göbl”.
- Puşcariu, Sextil. 1926. *Introducere. Gramatică. Caracterizarea dialectului istroromân*. Bucharest, Cultura Naţională, Institutul de Arte Grafice „Carol Göbl”.
- TC Cantemir, Traian. 1959. *Texte istroromâne*. Bucharest: Editura Academiei Populare Române.
- Kovačec A. Kovačec. 1984. “Istroromâna”. In *Tratat de dialectologie românească*, edited by Valeriu Rusu, 550-591. Craiova: Editura Scrisul Românesc.

## References

- Caragiu-Marioţeanu, Matilda, Ştefan Giosu, Liliana Ionescu-Ruxăndoiu, R. Todoran. 1977. *Dialectologie română*, Bucharest: Editura Didactică şi Pedagogică.



- Coteanu, Ion. 1957. *Cum moare o limbă – istroromâna*. Bucharest: Societatea de științe istorice și filologice.
- Geană, Ionuț. 2018. “Dialectul istroromân”. In *Convergențe balcano romanice în dialectele românești sud dunărene*, edited by Manuela Nevaci, Ionuț Geană, Carmen-Irina Floarea, and Teodora Toroipan, 111-129. Bucharest: Editura Universității din București
- Geană, Ionuț. 2017a. “On the Use of Compound Past in Istro-Romanian”. In *Sintaxa ca mod de a fi. Omagiu Gabrielei Pană Dindelegan, la aniversare*, edited by Adina Dragomirescu, Alexandru Nicolae, Camelia Stan, Rodica Zafiu, 207-212. Bucharest: Editura Universității din București.
- Geană, Ionuț. 2017b. “Voice in Istro-Romanian. A Corpus-Based Analysis”. *Fonetică și dialectologie*, XXXVI: 36-46.
- Maiorescu, Ioan. 1874. *Itinerar în Istria și vocabular istriano-român*, Iași: Tipo-litografia H. Goldner.
- Pescarini, Diego. 2018. “Stressed enclitics are not weak pronouns. A plea for allomorphy”. *Romance Languages and Linguistic Theory 14: Selected papers from the 46th Linguistic Symposium on Romance Languages (LSRL)*, 231-244. Stony Brook.
- Rusu, Valeriu (ed.). 1984. *Tratat de dialectologie românească*. Craiova: Scrisul Românesc.
- Sârbu, Richard, and Vasile Frățilă. 1998. *Dialectul istroromân*. Timișoara: Armacord.
- Vrzić, Zvezdana, and John Victor Singler. 2016. “Identity and language shift among Vlashki/Zheyanski speakers in Croatia”. In *Language Documentation and Conservation Special Publication No. 9 (January 2016): Language Documentation and Conservation in Europe*, edited by Vera Ferreira, and Peter Brouda, 51–68. Hawaii: University of Hawaii Press.
- Vrzić, Zvezdana. 2019. *Endangered Romance Languages in Istria, Croatia: Documentation and Preservation*. Abstract for LSRL49, University of Georgia (available URL: [http://linguistics.uga.edu/sites/default/files/LSRL49-abstracts/VRZIC\\_LSRL49.pdf?fbclid=IwAR2\\_5eCe9TIBRkKbCOMfhMlkg6wGSvWtZqYG3qwJ\\_9434KGHo8K8T3jd\\_cao](http://linguistics.uga.edu/sites/default/files/LSRL49-abstracts/VRZIC_LSRL49.pdf?fbclid=IwAR2_5eCe9TIBRkKbCOMfhMlkg6wGSvWtZqYG3qwJ_9434KGHo8K8T3jd_cao))
- Zegrean, Iulia. 2012. *Aspects on the Syntax of Istro-Romanian*. PhD Diss: Università Ca’Foscari, Venice, Italy.

## Notes

---

<sup>1</sup> For the purposes of this paper, by Easter Romance varieties I mean Daco-Romanian (DR), Aromanian (AR), Megleno-Romanian (MR) and Istro-Romanian (IR).

<sup>2</sup> For easy reading and comparison with standard Romanian, I chose to give the equivalent of IR into DR after the English translation.

# CHAPTER FIVE

## (BARE) OBJECTS OF PREPOSITIONS IN EASTERN ROMANCE

DANIELA ISAC

This paper focuses on definite DPs headed by null Ds in Eastern Romance (ER). Such DPs occur as objects of Ps in these languages and are theoretically challenging since null definite Ds are otherwise banned in Romance. The paper proposes that two factors contribute to the (c)overtness of the definite D in ER: (i) the definite article is affixal in these languages, which means that it never gets spelled out in the D head, but on some other head within the DP, and (ii) definite Ds undergo m-merge in ER, an operation that applies post-syntactically and reanalyzes two heads as one. In the proposed analysis D m-merges with the closest head it enters Agree with, regardless of the feature involved in Agree. The definite article is spelled out on this head only if D Agrees with it with respect to the [def] feature. Null definite articles are thus instances in which D m-merges with a head with which it Agrees in some other feature than [def], such as [Case] for instance.

Keywords: definiteness, null determiners, m-merge, Agree, prepositions

### 1. Introduction and Goals

Definite DPs in Eastern Romance (ER) languages - Romanian (Rom), Megleno-Romanian (MegR), Istro-Romanian (IR), Aromanian (Ar) have a peculiar property: the definite article can be null if these DPs are objects of prepositions that assign Accusative case, but not otherwise.

- (1) Şoricelul                    s-a        ascuns **de**    pisică. (Rom)  
little.mouse.def        REFL-has hidden from    cat  
'The little mouse hid from the cat.'

- (2) D. ieși într-un rînd **în piață**.  
 D. went.out in-one time in market  
 ‘One day D went to the market.’  
 (Ar, Cîndroveanu 1977:204)
- (3) ...și tîntea ăx căzú **di la măgar** un miğít.  
 ...and then CL.3P.S.DAT fell from at donkey a silver.coin  
 ‘... and a silver coin fell from the donkey’  
 (MegR, Saramandu et al 2016:177)
- (4) Mižulu căzut-a **dispre scînd**.  
 glass.def fallen-has off table  
 ‘The glass fell off the table.’  
 (IR, Zegrean 2012:178)

In this paper I will focus on simple objects of prepositions, i.e., objects of Ps that consist exclusively of a noun. A first goal of this paper is to provide an account for why this is the only type of DP in which the definite article can be null across ER. The proposal is that the affixal nature of the definite D in ER translates in definite Ds undergoing m-merge, an operation that reanalyzes two heads as one post-syntactically. More specifically, definite Ds m-merge with the closest head they enter Agree with and it is spelled out on this head only if the latter agrees with D in its definiteness ([def]) feature. With DPs objects of Ps the definite article can be null because in this case the closest head D agrees with is P, and P lacks a [def] feature.

The second goal will be to account for variation within ER with respect to whether the article must, or simply can be, null on objects of prepositions.

- (5) \*Femeia s-a uitat **către soare-le**. (Rom)  
 woman.DEF REFL-has looked towards sun.DEF  
 ‘The woman looked towards the sun.’
- (6) Muláreș si zăcătă și **cútru soar-li**  
 woman.DEF CL.REFL looked also towards sun.ACC.DEF  
 ‘The woman looked towards the sun as well.’

(MegR, Saramandu et al 2011:61)

- (7) io meg **la gospodar-u** (IR, Kovacek 1971:182)  
 I go to master.DEF  
 ‘I’m going to the master’
- (8) s-nu te-aspari ni **di sultan-lu**  
 SBJ-NEG CL.2P.SG.ACC-fear.2P.SG not.even of sultan.DEF  
 ‘Don’t be afraid even of the Sultan.’  
 (Ar, Candroveanu 1977:214)

As the above examples show, Romanian contrasts with all the other ER languages in that Romanian is the only one of these languages in which the definite article on objects of prepositions must be null. This paper accounts for this contrast by proposing that in all ER languages except Romanian the Num head may bear a contrastive ([c]) feature, which triggers the presence of a contrastive projection (ContrP) in the left periphery of the DP. NumPs headed by a Num head that bears a [c] feature raise to Spec,ContrP. The effect of the presence of ContrP is that P is no longer the closest head that D agrees with and that D no longer m-merges with P but with N, which is contained in the NumP that raises to Spec,ContrP.

The rest of the paper is organized as follows: in section 2, I will lay out the theoretical assumptions of the analysis, in section 3 I will present the proposal, section 4 will include an implementation of the proposal, case by case, and in section 5 I will present the conclusions.

## 2. Theoretical assumptions

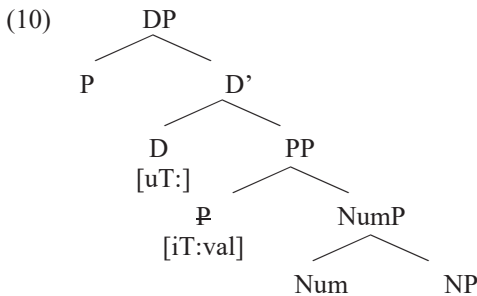
The analysis I will propose relies on the following important theoretical assumptions. First, I will adopt Pesetsky & Torrego's (2007) distinction between interpretable and uninterpretable features on the one hand, and valued and unvalued features on the other. Thus, I will assume a four way distinction between the following four types of features: (a) uninterpretable, valued features - [uF: val]; (b) interpretable, valued features - [iF: val]; (c) uninterpretable, unvalued features - [uF:]; (d) interpretable, unvalued features -[iF: ].

In applying this system to the features of the syntactic heads within DP, I will follow Cornilescu and Nicolae (2009, 2011) in assuming that Romanian nouns bear a valued [def] feature and that other items within the DP that have a [def] feature have an unvalued instance of this feature. This view will be extended to all the other ER languages.

- (9) D<sub>[i:def: ]</sub> N<sub>[u:def:+]</sub>

Given its unvalued [def] feature, D searches for a matching feature in its c-command domain and it finds the [def] feature on N. Agree takes place, and the value provided by N will then be shared by D. In what follows I will call the set of items that enter Agree and come to share the value of a given feature a 'valuation chain'. Thus, the definiteness valuation chain in (9) includes D and N.

Apart from an unvalued [def] feature, D also bears a Case feature. I will follow Pesetsky and Torrego (2004) in assuming that all structural cases are instances of [uT] on D and that objects of prepositions get their [uT], i.e. Case, feature valued by the preposition. Moreover, I will also assume, together with Pesetsky and Torrego (2004) that prepositions are merged DP internally, in a position that is analogous to the position occupied by T within the CP. More specifically, P is lower than the DP (the analogous of CP in the clausal domain) and higher than NumP (the analogous of the vP phase in the clausal domain). When D searches for a matching [T] feature, it finds P and agrees with its [T] feature. The value of the [T] feature on P is thus shared with D. In other words, D initiates a second valuation chain, apart from the definiteness one, i.e., the [T] chain, which contains D and P. Last but not least, Pesetsky and Torrego (2001, 2004, 2007) propose that Ps raise to D to check the latter's EPP feature. Given that in Pesetsky and Torrego's (2004) view head movement and phrasal movement both target the specifier position of the attracting head, when the P head is attracted by D, it moves to Spec,D, as shown in (10).



Even though head movement and phrasal movement target the same position, the choice between the two is not random, but depends on how local the goal is.

(11) Pesetsky and Torrego (2001, 2007):

if a head H attracts a feature of XP as part of a movement operation, then

- a. if XP is the complement of H, **copy the head** of XP into the local domain of H
- b. otherwise, **copy XP** into the local domain of H

Thus, if a projection intervenes between D and the PP, it is the PP that moves to Spec,D, rather than the P head alone. This will be an important difference in our analysis.

### 3. The proposal

#### 3.1 Previous literature

The existing generative literature on definiteness in ER languages is predominantly focused on Romanian. Moreover, very little attention has been given to definite objects of prepositions with null articles in Romanian or any of the ER languages. The one exception is Mardale (2006), who proposes that definite D is null with objects of prepositions because definite D incorporates into P, a process which is made possible by the reduced syntactic structure of these DPs. In particular, objects of prepositions are assumed to lack a K(ase)P, which would normally be projected in between PP and DP, as shown in (12).

(12) [PP P [(KP) (K) [DP D [NumP Num [NP N ] ] ] ] ] ]

Moreover, the reason why D must incorporate into P, as opposed to being pronounced on N, is that the structure of objects of Ps is defective in yet another way, according to Mardale (2006). In particular, objects of Ps also lack a NumP. Since Mardale (2006) assumes Dobrovie-Sorin and Giurgea's (2006) analysis of definiteness in Romanian, according to which D lowers to Num and gets spelled out on the N which has raised to Num, the absence of Num accounts for the fact that D will not get spelled out on N, but will incorporate into P.

Mardale's (2006) analysis raises a number of questions. First, in Dobrovie-Sorin and Giurgea's (2006) view, D-lowering targets the head of D's complement. In the absence of NumP, the complement of D is NP, so the prediction should be that D will lower to N in this case, contrary to Mardale's (2006) analysis. Second, Mardale's (2006) analysis cannot

straightforwardly be extended to the other ER languages, in which the definite article can be null on objects of prepositions, but can also be overt.

### 3.2 Current Proposal

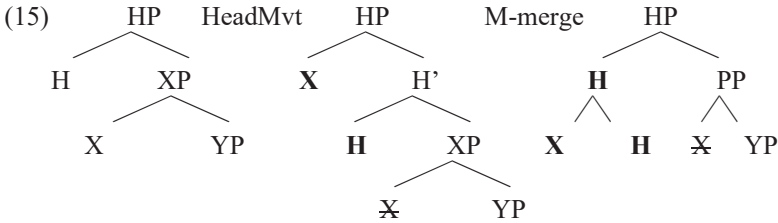
In order to account for the data presented in section 1, I propose the following two rules apply in ER.

- (13) *M-merge*:  
Definite D m-merges with the closest head it Agrees with.
- (14) *Spell-Out*:  
The definite article is spelled out on the head that definite D m-merges with iff that head bears a [def] feature. Otherwise, the definite article is phonologically null.

Several comments are in order with respect to these two rules. First, I am adopting a general definition of m-merge as an operation that reanalyzes two heads as one in the morphological component. This is similar to the notion of 'Morphological Merger' proposed by Marantz (1984), as well as to the 'M-merge' operation proposed in Matushansky (2006), but there are important differences to note. The differences concern mainly the input to the merger operation. In Marantz's (1984) view, Morphological Merger applies to any two items X and Y that are in a given relation. The two items that can undergo Morphological Merger can be phrases or heads and the nature of the relation between X and Y is not made explicit. In contrast, the input to the m-merge operation proposed in this article consists of two heads in a particular relationship: Agree. More specifically, m-merge operates on all the Agree chains initiated by D--the definiteness chain, which includes D and N, and the Case chain, which includes D and P. The head that is closest to D across both of these chains will undergo m-merge with D.

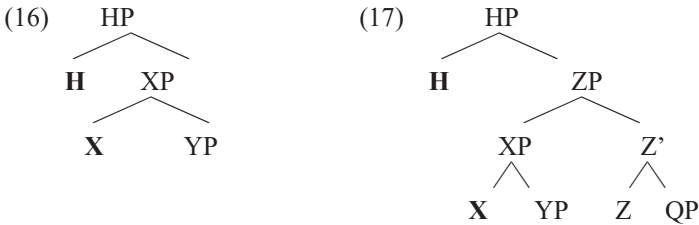
The m-merge operation proposed here also differs from Matushansky's (2006) M-merge operation. In Matushansky's view the input for M-merge consists of two heads in a Spec-head configuration. This configuration is in itself the result of head movement, which, as discussed above, targets the Spec position of an attracting head, just like phrasal movement.





Thus, in (15), X moves as a head to the specifier of the attracting head Spec,H and further undergoes M-merge with H.

In contrast, in the view adopted here, there is no restriction on the type of configuration that constitutes the input to the m-merge rule. In other words, not only heads that are in a Spec-head configuration can undergo m-merge, but a head H may m-merge with another head X when X is the head of H's complement for example, or when X is the head of a phrase in the Spec of H's complement, as represented in (16) and (17).



Thus, in the analysis proposed here, it is a consistent property of definite Ds in ER to always undergo m-merge, whereas in Matushansky's (2006) view M-merge only affects Ds that attract a head in their Spec.

The rule in (13) is also similar in spirit to Dobrovie-Sorin and Giurgea's (2006) lowering rule. In their view definite Ds are subject to a PF rule which lowers the D head to the head of its complement, i.e., Num. Just as our m-merge rule in (13), Dobrovie Sorin and Giurgea's (2006) lowering rule is post-syntactic. However, our m-merge rule differs from their proposal in several respects. Unlike their D lowering rule, which always targets the head of D's complement, the rule in (13) is based on the linear order of heads at PF, rather than on syntactic structure. This allows Ds to m-merge with the head of an XP in the Spec of its complement, for example, rather than strictly with the head of its complement. Second, our m-merge rule differs from their proposal in the way in which it feeds the spell-out of the definite article. In Dobrovie-Sorin and Giurgea's proposal D ends up pronounced as a suffix on N because of two factors: one is that D lowers to the head of its

complement, which is Num, and the other one is that N independently raises to Num in the syntax. Since D and N always meet in Num, D always gets pronounced on N in simple (i.e., unmodified) DPs in their view. In contrast, in our view, the spell-out rule applies to a subset of the heads that are the output of our m-merge rule to the output of the m-merge rule, i.e., only to those heads that participate in the valuation of the [def] feature. While D can m-merge either with a head in its definiteness valuation chain or with a head in its Case, i.e., T chain, the spell out rule allows only heads that are part of the definiteness valuation chain to host the overt definite article.

Finally, the rule in (14) is also similar to Cornilescu and Nicolae's (2011) proposal that the definite article is pronounced at PF on the highest item below D bearing a [def] feature, on condition this item has nominal features (i.e., is a noun or an adjective). The differences between the rule in (14) and this proposal have to do with m-merge, an operation that is part of the current proposal but not Cornilescu and Nicolae's. What m-merge achieves for the analysis in this paper is the possibility of accounting for the (lack of) spell out of definite articles with objects of prepositions.

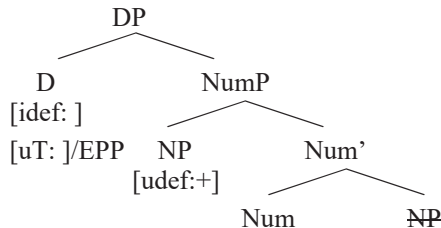
## 4. Implementation

In this section I will show how the above proposal can account for the realization of the definite article on simple DPs in ER, i.e., DPs that contain a noun only.

### 4.1 Definite DPs in non-prepositional environments

The definite article of definite DPs in non-prepositional environments is always expressed overtly as a suffix on the noun in all ER languages. In order to account for this, I will rely on assumptions made by Cornilescu and Nicolae (2009, 2011), Tănase-Dogaru (2012) a.o. for Romanian, according to which DPs contain a NumP and that the Num head attracts the NP to its Spec. This assumption will be generalized to all ER languages<sup>1</sup>.

(18)



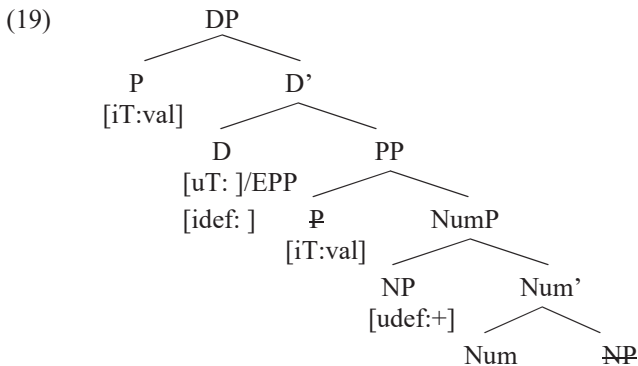
Recall that D bears two unvalued features in our analysis - [def] and [T]. According to the m-merge rule in (13), D will m-merge with the closest head that it establishes an Agree relation with. Given that D has two unvalued features, there are two potential Agree relations that D could establish. When D searches for a goal with a matching valued [def] feature, it finds N. Hence D enters Agree with N. On the other hand, when D probes within its c-command domain for a matching [T] feature, it finds no match. D will nevertheless remain active in the derivation for future Agree relations, since D is the head of the DP phase. The [uT] feature on D is thus checked and valued from outside of the DP. What is important is that there is no [T] chain within the DP and that the only Agree relation established by D within the DP domain is with the N head. Thus, D will m-merge with N according to the m-merge rule in (13) and will be spelled out on N according to the rule in (14).

## 4.2 Definite objects of Ps

### 4.2.1 Bare definite objects of Ps

Definite objects of Ps can be bare across ER, as illustrated in (1)-(4). The goal of this section is to account for why the definite article can be covert with objects of Ps but not with other definite DPs.

As discussed above, we assume, together with Pesetsky and Torrego (2004) that Ps are merged inside the DP, and then move to Spec,D.



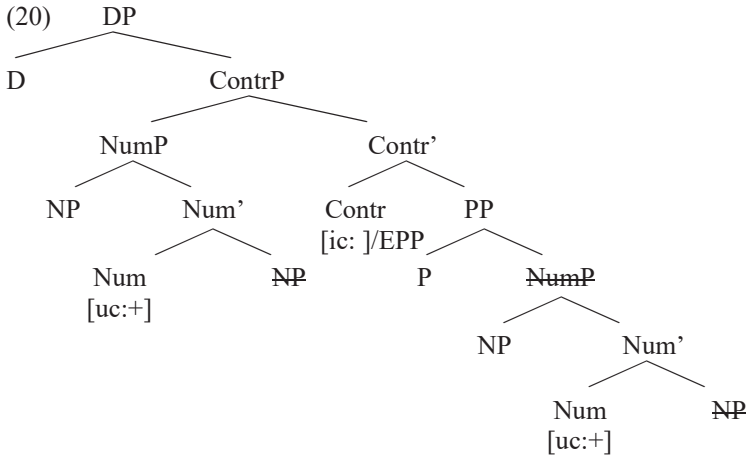
D initiates two Agree chains within the DP: one with N, for the [def] feature and one with P, for the [T] feature. Given that P is closer to D than N, D will undergo m-merge with P, according to the rule in (13). Moreover,

given that P has no [def] feature, the definite article will remain phonologically null, according to the Spell-out rule in (14). Thus, the reason why the definite article can be covert with objects of Ps but not with other definite DPs in ER is that P offers a valued match for the unvalued [T] feature on D and is thus a candidate for m-merge with D, while in the absence of P, the only head that D Agrees with inside the DP is N and thus D m-merges with N.

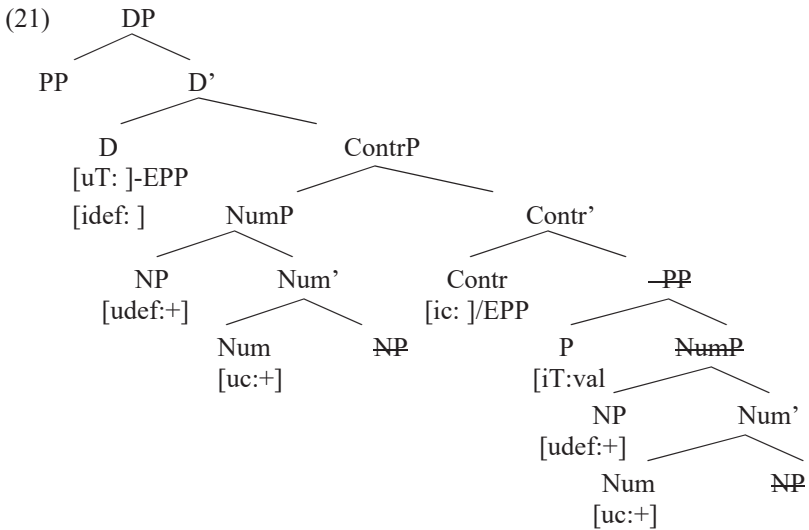
#### 4.2.2 Overtly definite Objects of Ps

As discussed in section 1, in some ER languages the object of P does not have to be bare when definite and can also be overtly definite. This is an option in all ER, with the exception of Romanian, where unmodified definite objects of Ps are always bare. This raises the following three questions: (i) how can we account for overtly definite objects of Ps in the ER languages that allow such objects? (ii) what is the difference between bare definite objects of Ps and overtly definite objects of Ps? and (iii) how can we explain the fact that Romanian does not allow overtly definite objects of Ps? This section will address the first question, while the other two will be the topic of sections 4.2.3 and 4.2.4, respectively.

In order to offer an account of overtly definite objects of Ps in the ER languages that allow such objects, I will propose that objects of Ps with overt definite articles include a periphery projection that I will call *Contr(astive)P*, for reasons I will detail below. The presence or absence of projections in the left periphery of the DP is pre-determined by the features of the items in the numeration. If a sub-constituent within the DP bears an uninterpretable [c] feature for example, the DP will have to contain a head with a matching [c] feature in its left periphery, i.e. a *Contr* head. I propose that the *Num* head may contain a [c] feature in all ER languages except Romanian. The structure I propose for overtly definite objects of Ps in ER is thus the following:



Assuming that the [c] feature on the Contr head is associated with an EPP feature, the Contr head will always attract a constituent with a matching [c] feature to its Spec. Thus, in (20), the NumP phrase raises to Spec,Contr to check the EPP associated with the [c] feature on the Contr head. One consequence of the presence of a ContrP in the structure is that D will not be able to check the EPP on its [T] feature by attracting the P head. This is because the PP is no longer the complement of D when the ContrP is present. Instead, it is the whole PP, or rather the remnant PP, that raises to Spec,D.



The other consequence of the presence of a ContrP, one that is most important for our analysis, is that the heads that D enters Agree with are now in a different configuration. In particular, in a structure like (21), the N head, i.e. the only head in the definiteness valuation chain, is closer to D than P, the only head in the [T] valuation chain. Hence, D will m-merge with N according to the rule in (13) and will be spelled out on N according to the rule in (14).

#### 4.2.3 Bare vs overtly definite objects of Ps in ER

I will now address the second question stated above, i.e. what is the difference between bare definite objects of Ps and overtly definite objects of Ps in ER languages that allow both? From a morpho-syntactic point of view, the difference is obviously related to the features of the Num head. In ER languages that allow both overtly definite objects of Ps and bare ones the Num head may bear a [c] feature or not. If Num does have a [c] feature, the structure is as in (21) and the definite article is overtly expressed on the object of P, whereas if the Num head lacks a [c] feature, the proposed structure is as in (19) and the definite article is null.

In order to identify the semantic difference between bare definite objects of Ps and overtly definite objects of Ps in ER languages, I will take Hawkins' (1978) classification of the uses of the definite article as a starting point.

According to Hawkins (1978), definite expressions can be used in the following situations:

(i) to refer back to the referent of an antecedent expression (the ‘anaphoric use’), as in (22), where the definite DP *the dog* is anaphorically co-indexed with the previously introduced expression *a dog*.

(22) A man showed up with *a dog*<sub>i</sub>. *The dog*<sub>i</sub> looked anxious.

(ii) to refer to individuals which are present in the speech situation and uniquely meet the descriptive content of the definite description in that situation (the ‘immediate situation’ use). For example, in a household that includes only one dog, the definite *the dog* refers to the unique dog in that situation.

(23) Did you take *the dog* out?

(iii) to refer to individuals that uniquely meet the descriptive content of the definite description, even if they are not present in the immediate utterance situation. Rather, the referent of the definite description is part of a larger situation, such as the speaker’s neighbourhood, or the speaker’s country or town (the ‘larger situation’ use). For example, the definite expression ‘the Prime Minister’ refers to a unique individual whose identity depends on the country that the utterance situation is part of.

(iv) in order to relate back to a contextually present expression in an indirect way (the ‘associative anaphoric’ or ‘bridging’ use). This is a special case of the anaphoric use, in the sense that the definite is not co-referential with the antecedent, but stands in some salient relationship to it.

(24) John bought *a book* today. *The author* is French.

The entity referred to by the definite expression is strictly speaking new to the discourse, but its existence is linked to the old, familiar referent of *a book*. In this particular case the relation between the two referents is one of producer-produced.

A first hypothesis to consider is whether the two types of definite objects of Ps in ER are semantically specialized in that they each cover a particular subset of the uses included in Hawkins’ (1978) classification. What the examples below show however is that this hypothesis has to be rejected, since one and the same use in Hawkins’ (1978) classification can be

expressed either by a definite bare object of P or by an overtly definite one. In (6), from Megleno-Romanian, for example, the referent of an overtly definite object of P - *soarli* 'the sun' is an instance of the 'immediate situation' use. However, the same referent, in the same situation ('immediate situation' use) can also be expressed by a bare definite object in Megleno-Romanian.

- (25) Dăscărcă lúcrili și cățo la spilárea.  
 unloaded.3P.SG clothe.DEF.PL and started.3P.SG at washing.DEF  
 Ca dăsfărșó, li tinsi **la soari**  
 When finished.3P.SG, CL.3P.PL.ACC hung.3P.SG at sun  
 s-úscă.  
 CL.REFL-dry.3PL  
 '(S)he unloaded the clothes and proceeded with their washing. When she finished, she hung them up in the sun, to dry.'  
 (MegR, Saramandu et al 2016:195)

I will propose instead that the semantic difference between bare definite objects of Ps and overtly definite ones has to do with notions such as familiarity vs uniqueness, which are at the core of the two most influential theories about definiteness.

Familiarity based theories argue that definite descriptions refer to individuals that are in some sense familiar to the discourse participants (Heim 1982, 1983, Roberts 2003, etc.). There are several ways in which an individual can be familiar: by being linked to an old referent (i.e. either by being anaphoric to a preceding linguistic expression or by being somehow associated with an old discourse referent), by being perceptually accessible to the discourse participants, or by being 'globally familiar in the general culture' (Roberts 2003:304).

Under uniqueness-based theories, on the other hand, definite descriptions are used to refer to things that can be uniquely identified in a particular context or situation (Russell 1905, Lewis 1979, Kadmon 1990, Gundel et al 1993, etc.).

Both of these theories claim to be able to account for all the uses of definite descriptions, i.e., for all the uses in Hawkins' (1978) classification. However, in spite of this overlap, there is evidence that uniqueness and familiarity play distinct roles in the analysis of definiteness and thus that both theories are needed. Schwartz (2009) for example shows that standard German has two types of definite articles: a 'weak' one, which encodes uniqueness, and a 'strong' one, which is anaphoric in nature. I propose that ER languages offer additional evidence that both uniqueness and familiarity



are needed in the analysis of definiteness in natural languages. More specifically, I propose that in ER languages that allow objects of Ps to be either bare or overtly definite, bare objects of P express familiarity, while objects of Ps with an overt definite D express uniqueness. Thus, in (6), which contains an overtly definite object of P in an ‘immediate situation’ use, the speaker refers to ‘the sun’ as the unique entity that is present in the speech situation and meets the descriptive content of the definite description (the sun as a unique entity), whereas in (25), which contains a bare definite object of P in an ‘immediate situation’ use, the speaker refers to the sun as an entity that is familiar to the discourse participants, given that it is perceptually accessible to the discourse participants (recall that this is one way in which an individual counts as being familiar to the discourse participants).

Two types of arguments can be made in defense of this proposal. The first argument has to do with the correlation between the overt definite D and the existence of ContrP. Recall that in section 4.2.2 we proposed that objects of P headed by an overt D always contain a ContrP in ER. ContrP is placed in between the DP and the PP in the hierarchy of projections and this prevents P from raising as a head to Spec,D. As a result, D m-merges and gets spelled out on a head within the DP. What is important is that there is a correlation between the presence of ContrP and the overtness of the definite article. Given this link, it is not surprising that the semantics of the Contr head will contribute to the interpretation of a DP headed by an overt definite D. Under the assumption that the Contr head introduces a set of alternatives (Rooth 1992, Krifka 1993, etc.) and a contrast between the referent of the constituent that bears the [c] feature and the other alternatives, the Contr head guarantees ease of identifiability for the referent that is set in contrast with the other alternatives. More specifically, objects of P that contain a ContrP and are interpreted as contrastive in the sense of uniquely identifying a referent present in the immediate or larger situation, to the exclusion of other possible referents that are not uniquely identifiable. Hence the argument that can be made is that the presence of ContrP is conducive to a uniqueness interpretation for the definite D. Given that the presence of Contr also leads to an overt definite D, a link is established between definite objects of P with overt Ds and a uniqueness interpretation.

A second argument in support of the proposal that overt definite Ds are correlated with a uniqueness interpretation has to do with objects of Ps that are modified, or contain a complement. Modifiers and complements add properties and as such increase the identifiability of the referent. If our analysis is on the right track, we expect definite modified objects of Ps, or objects of Ps that contain complements to always contain an overt definite

D. This expectation is borne out in all ER languages. For lack of space, I have included only an example from Megleno-Romanian.

- (26) *ân mârđina di váli* (MegR, Atanasov 2002:275)  
 in edge of river  
 ‘on the river side’

I will conclude that all ER languages with the exception of Romanian (which will be discussed below) definite objects of Ps can be either bare, if they express familiarity, or headed by an overtly definite D, if they express uniqueness.

One more clarification needs to be made before we move on. We argued above that ER languages offer evidence that definite objects of Ps are interpreted differently depending on whether they are headed by an overt definite D or a covert one. This pattern is similar to what has been observed in standard German (Schwartz 2009) where familiarity and uniqueness are expressed by two different forms of the definite article (the strong one and the weak one). In spite of some similarity though, there are important differences between ER and German. In particular, while in German the difference between the two interpretations is encoded in the article itself (there are two different lexical Ds, with different semantic interpretations), in ER the semantic difference between objects of Ps headed by a null definite D and those headed by an overt definite D is the outcome of two factors: (i) the features of the Num head, rather than the D head (in particular whether Num bears a [c] feature or not); and (ii) the presence or absence of a P. The two different definite Ds (null vs overt) are not different lexical items in ER and the overtness of the definite D is the result of a post syntactic Spell-Out rule whose application depends on these two factors.

Further evidence that ER languages do not have two different definite Ds is offered by DPs that occur in non-prepositional environments. If the two Ds (the overt D and the covert one) were independent lexical items in ER, we would expect regular DPs to be possibly headed either by the overt D or by the null D, contrary to fact. All definite DPs that are not objects of Ps must be headed by an overt D in ER. The interpretation of these DPs can be both familiar and unique, depending on whether the Num head bears a [c] feature or not, and hence depending on the presence or absence of a ContrP in the periphery of the DP. However, in both cases the definite D is spelled out overtly. This is because in both cases there is no P and the only valuation chain established by D is the definiteness valuation chain which contains the N. The fact that the definite D is always spelled out as overt in these cases shows that the difference between familiarity and uniqueness

interpretations of these DPs does not follow from the properties of D, but from the application of the Spell-out rule.

#### 4.2.4 Romanian vs other ER languages

I will now address the third question mentioned at the beginning of section 4.2.2: how can we explain the fact that Romanian does not allow overtly definite objects of Ps, in contrast with all the other ER languages? In particular, Romanian definite objects of Ps are always bare if unmodified, regardless of whether they receive a familiarity interpretation or whether they are interpreted as unique in a situation.

(27a) illustrates the familiarity meaning of definite bare objects of P in Romanian: the object of the preposition *pe* ‘on’ is coreferential with the previously mentioned DP *o canapea* / ‘a couch’ and is thus an instance of an anaphoric use. The contrast below shows that definite objects of Ps that refer to a familiar object must be bare.

- (27) In sufragerie erau **o canapea** și două fotolii.  
 in living.room were a couch and two armchairs  
 ‘In the living there were a couch and two armchairs.’
- a. **Pe canapea** dormea o pisică. (Rom)  
 On couch sleep.3P.SG.IMP a cat.  
 ‘There was a cat sleeping on the couch.’
- b. \***Pe canapea-ua** dormea o pisică. (Rom)  
 On couch.DEF sleep.3P.SG.IMP a cat.  
 ‘There was a cat sleeping on the couch.’

In (28) on the other hand, the object of P denotes a uniquely identifiable object, and the definite article must be again covert.

- (28) Maria se uită **către soare** / \***către soare-le**. (Rom)  
 Maria CL.REFL looked towards sun / towards sun.DEF  
 ‘Maria looked towards the sun.’

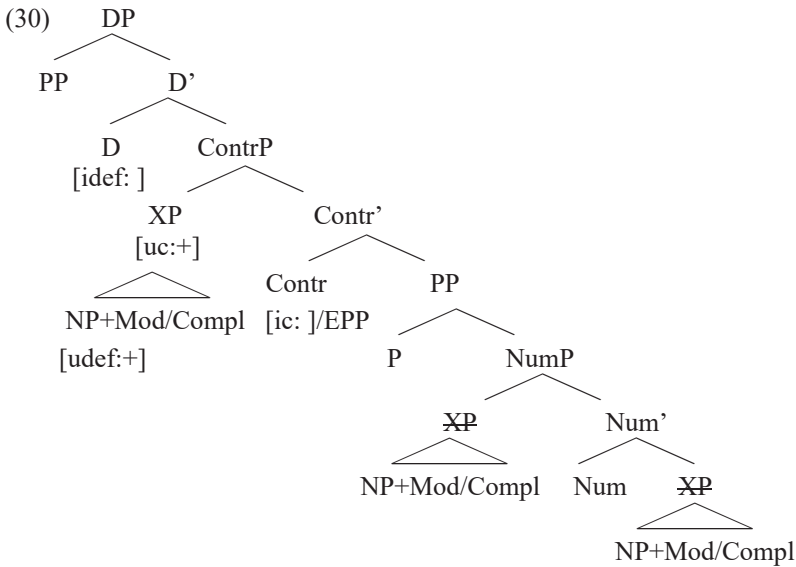
Recall that in our analysis of ER languages other than Romanian we proposed that the overt definite D on an object of P signals uniqueness. The fact that Romanian objects of Ps are never overtly definite suggests that Romanian does not make a formal distinction between definites that express familiarity and those that express uniqueness, and that these interpretations can be obtained by contextual entailments in Romanian. However, unlike objects of P that are reduced to a noun, definite objects of Ps that contain a

modifier or a complement must always be overtly definite in Romanian, just as in the other ER languages.

- (29) în centru\*(-I) vechi (Rom)  
 in center.DEF old  
 ‘in the old center/town’

Under the assumption that modified Ns or Ns with a complement are contrastive just as much as their counterpart in the other ER, at least in these cases there must be a ContrP in the structure of Romanian DPs, and hence at least in these cases the uniqueness interpretation springs from a morpho-syntactic feature.

The question now is why the presence of a modifier/complement in Romanian enforces a ContrP and thus a uniqueness interpretation. It is important to notice that the presence of ContrP depends on the presence of a [c] feature on one of the nominal sub-constituents. This is because in our analysis the Contr head bears an interpretable but unvalued [c] feature and as such it needs to probe for a matching valued feature in its c-command domain. If no match is found, the [c] feature on the Contr head remains unvalued and the derivation would crash. The key is that a matching valued [c] feature can be offered either by the Num head or by some other constituent, like the XP that contains the modifier/complement. If it is the XP containing the modifier/complement that bears a [c] feature, the representation of the DP is as in (30). The XP marked as [c] is attracted to Spec,Contr to check the EPP on the [c] feature in Contr. The N bearing the valued [def] feature is part of this XP and is carried along to Spec,Contr, i.e. to a position contiguous to D. Thus, when D searches for a valued match for its [def] feature, it finds N, and when D searches for a match that could provide a value for its [T] feature, it finds P. Two valuation chains are initiated by D within the DP, but the closest item that enters Agree with D is N. Hence, D m-merges with N according to the rule in (13). Moreover, the definite article is also spelled out on N, given that N participates in the definiteness valuation chain.



The difference between Romanian and the other Eastern Romance languages has to do with the nominal sub-constituents that bear a [c] feature. In Romanian only constituents that license modifiers or complements can bear such a feature, whereas in the other Eastern Romance languages, the Num head can also bear a [c] feature, in addition to constituents that license modifiers or complements. The effect of this difference is that in Romanian only definite objects of Ps that contain a modifier or a complement are overtly definite, whereas in the other Eastern Romance languages this is the case not only for definite objects of Ps that contain a modifier or a complement, but also for unmodified objects of Ps that contain a Num marked as [c].

## 5. Conclusions

This paper focused on a peculiar property of definite objects of prepositions in Eastern Romance: these objects can be bare in spite of the fact that their interpretation is clearly definite and in spite of the fact that the definite article is otherwise overt in these languages. Moreover, while in Romanian unmodified objects of Ps must be bare, in the other Eastern Romance languages the definite article could be either null or overt.

In order to account for these properties, we proposed that definite Ds in Eastern Romance must undergo m-merge, an operation that analyzes two

heads as one in the morphological component. The exact head the D m-merges with depends on the Agree chains initiated by definite D within the DP. More specifically, D m-merges with the closest head it enters an Agree relation with. The overtness of the definite D depends in our analysis on the application of an additional rule, Spell-out. This rule spells out the definite article on the head that D m-merges with on condition this head bears a [def] feature. Thus, only a subset of the heads that D can possibly m-merge with are possible hosts for the overt definite article in our view, i.e., only those heads that are part of the definiteness valuation chain. Thus, the definite article on objects of Ps in Eastern Romance can be null in cases in which D m-merges with a head that does not bear a [def] feature, i.e., when it m-merges with P. In order to account for the differences between Romanian and the other Eastern Romance languages we proposed that these differences spring from the features of the Num head in these languages. In particular, the Num head in all Eastern Romance languages except Romanian may bear a [c] feature, which triggers the presence of a ContrP in the DP. Given that the ContrP is placed between the DP and the PP, the effect of the presence of a ContrP is that P is no longer the closest head that D agrees with and hence D m-merges with N rather than with D. Hence in these cases the definite article is spelled out overtly on N according to our analysis. Since Num cannot bear a [c] feature in Romanian, unmodified objects of Ps in this language are never overtly definite.

We also proposed that the presence of a ContrP has semantic consequences. In particular, the presence of a ContrP leads to an overt definite D and to an interpretation of the DP in terms of uniqueness, whereas the absence of a ContrP is correlated to a familiarity interpretation and to a covert definite D. Eastern Romance languages thus offer evidence that uniqueness and familiarity play distinct roles in the grammar, as the two different phonological realizations of the definite D are mapped to two different semantic interpretations.

## References

- Atanasov, Petar. 2002. *Meglenoromâna astăzi*. București: Academia Română.
- Cândroveanu, Hristu. 1977. *Antologie de proză aromână*. București: Editura Univers.
- Cornilescu, Alexandra and Alexandru Nicolae 2009. "On the syntax of Romanian definite phrases. Changes in the patterns of definiteness checking". In *The Noun Phrase in Romance and Germanic: Structure,*

- variation, and change*, edited by Petra Sleeman, Harry Perridon, 193-221. Amsterdam: John Benjamins.
- Cornilescu, Alexandra and Alexandru Nicolae 2011. "Nominal Peripheries and phase structure in the Romanian DP". *Revue Roumaine de Linguistique* LVI (1): 35–68.
- Dobrovie-Sorin, Carmen and Ion Giurgea 2006. "The Suffixation of Definite Articles in Balkan Languages". *Revue roumaine de linguistique* LI (1): 73-104.
- Gundel, Jeanette, Nancy Hedberg and Ron Zacharski. 1993. "Cognitive status and the form of referring expressions in discourse". *Language* 69 (2): 274 -307.
- Hawkins, John A. 1978. *Definiteness and Indefiniteness: a Study in Reference and Grammaticality Prediction*. London: Croom Helm and Atlantic Highlands NJ: Humanities Press.
- Heim, Irene. 1982. "The semantics of definite and indefinite noun phrases." PhD diss., University of Massachusetts.
- Heim, Irene. 1983. "File change semantics and the familiarity theory of definiteness". In *Meaning use and interpretation of language*, edited by Rainer Bäuerle, Christoph Schwarze and Arnim von Stechow, 164-189. Berlin, Boston: Walter de Gruyter.
- Kadmon, N. 1990. "Uniqueness." *Linguistics and Philosophy*. 13 (3): 273-324.
- Kovacek, August. 1971. *Descrierea Istroromânei actuale*. București: Editura academiiei.
- Krifka, Manfred. 1993. "Focus and Presupposition in Dynamic Interpretation". *Journal of Semantics* 10 (4): 269-300.
- Lewis, D. 1979. "Score keeping in a language game". *Journal of Philosophical Logic*. 8 (1): 339-359.
- Mardale, Alexandru. 2006. "Why *on table* is *on the table*? Notes on the definite article's incorporation with prepositions in Romanian". In *Proceedings of the First Central Europe Student Conference in Linguistics* edited by Beáta Gyuris, 1-12. Budapest: Research Institute for Linguistics.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, Mass.: MIT Press.
- Matushansky, Ora. 2006. "Head movement in linguistic theory". *Linguistic Inquiry* 37 (1): 69–109.
- Pesetsky, David and Esther Torrego 2001. "T-to-C Movement: Causes and Consequences". *Ken Hale: A Life in Language*, edited by Michael Kenstowicz, 355-426. Cambridge, Mass.: MIT Press.

- Pesetsky, David and Esther Torrego. 2004. "Tense, Case and the Nature of Syntactic Categories". *The Syntax of Time*, edited by Jacqueline Guéron and Jacqueline Lecarme, 495-538. Cambridge, Mass.: MIT Press.
- Pesetsky, David and Esther Torrego. 2007. "The Syntax of Valuation and the Interpretability of Features". *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation. Linguistik Aktuell/Linguistics Today 101*, edited by Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262-294. Amsterdam: John Benjamins.
- Roberts, Craige. 2003. "Uniqueness in definite noun phrases". *Linguistics and Philosophy* 26 (3): 287-350.
- Rooth, Mats. 1992. "A Theory of Focus Interpretation". *Natural Language Semantics* 1 (1): 75-117.
- Russell, Bertrand. 1905. "On denoting". *Mind* 14 (56): 479-93.
- Saramandu Nicolae, Marilena Tiugan, Irina Floarea and Alina Celac. 2011. "Dicționar meglenoromân". *Fonetică și dialectologie*, XXX: 51-115.
- Saramandu Nicolae, Marilena Tiugan, Irina Floarea & Alina Celac. 2016. "Dicționar meglenoromân". *Fonetică și dialectologie*, XXXV: 163-204.
- Schwartz, F. 2009. *Two Types of Definites in Natural Language*. PhD diss. University of Massachusetts Amherst.
- Tănase-Dogaru, Mihaela. 2012. *The Syntax of Quantity and Quality in Romanian*. Bucharest: Editura Universității din București.
- Zegrean, Iulia-Georgiana. 2012. *Balkan Romance: Aspects on the Syntax of Istro-Romanian*. PhD diss., Università Ca'Foscari Venezia.

## Note

---

<sup>1</sup> There are other functional heads in the NumP, but for the purposes of this paper I will assume the simplified version in (23)



CHAPTER SIX

THE (UN)BOUNDEDNESS  
OF QUANTIFIER SCOPE:  
EVIDENCE FROM HUNGARIAN

GERGŐ TURI AND BALÁZS SURÁNYI

According to the classic generative view, a quantifier cannot take scope outside of the finite clause in which it is located. The classic approach assumes that this constraint is grammatical in nature, and it derives from the locality properties of the covert movement operation of Quantifier Raising (QR, May 1977, 1985). Some apparent exceptions to the finite clause-boundedness restriction have been noted in the literature in which the scope of a quantifier apparently extends beyond the subjunctive finite clause in which it is located (it has *extra wide* scope). The present experimental study, conducted in Hungarian, investigates whether extra wide scope-taking from subjunctive complement clauses is indeed systematically available, and if so, whether it may be licensed by the syntactic transparency of subjunctive complement clauses, a position developed in Wurmbrand (2013, 2018), or by a process of semantic reanalysis, as proposed in Farkas and Giannakidou (1996). Based on ratings of wide scope out of non-finite clauses, simple subjunctive clauses, and subjunctive clauses within complex NPs (each type functioning as a complement to a transitive verb), we argue that while extra wide scope out of subjunctive complement clauses is indeed systematically licensed, the overall pattern of results can be given a more uniform account by semantic reanalysis than by excepting subjunctive complement clauses from finite clausal bounding domains.

Keywords: quantifier scope, Quantifier Raising, locality, syntactic island, subjunctive, semantic reanalysis, Hungarian

## 1. Introduction

Sentences containing multiple quantifier phrases (QP) are often scopally ambiguous in natural languages. (1) below is ambiguous between a surface scope (1a) and an inverse scope (1b) reading.

- (1) [[Exactly two students] [passed every test]].
- a. Exactly two students are such that they passed every test.
  - b. Every test is such that exactly two students passed it.

Mainstream generative theories analyze similar inverse wide scope readings as being the result of a movement transformation (May 1977, 1985). The supposed movement, Quantifier Raising (QR), covertly raises the quantifier expression from its surface position to its scope taking position without changing its place in the uttered linearization, resulting in a covert syntactic representation along the lines of (2).

- (2) [ [Every test x]: [[exactly two students] [passed x]]

QR as a movement operation exhibits similarities to well-known overt movement operations such as *wh*-movement. A fundamental syntactic parallel is that both operations can only raise the moved element to a position from which it c-commands its original position. Another basic similarity is that both QR and *wh*-movement give rise to Weak Crossover effects (Chomsky 1976). Further, neither QR nor *wh*-movement can raise any element out of strong islands (Ross 1967).

In the face of these parallels between *wh*-movement and QR, the two have been assumed to be distinct when it comes to boundedness. *Wh*-movement is known to be able to move the *wh*-element to a position in a clause that is superordinate to the finite clause in which the *wh*-element is base generated (3a) ('long' *wh*-movement). On the other hand, it has been assumed that QR cannot be long in the same way as *wh*-movement (and other A-bar movements, including focus movement, topicalization and relativization) can. Thus, the universal QP in (3b) cannot be interpreted with inverse wide scope over the matrix indefinite, a reading paraphrased in (3c). This has been taken to show that QR cannot be long: it is bounded by the smallest finite clause that contains the QP in base structure (see Farkas 1981; Fodor and Sag 1982; May 1985; Beghelli 1993; Fox and Sauerland 1996; Szabolcsi 1997); see (4).

- (3) a. Which student did you say [ that you will meet \_\_\_]?  
 b. A girl said [ that each boy had met you].  
 c. For each boy there is a potentially different girl who said that he had met you.
- (4) *Classic finite clause boundedness constraint on quantifier scope*  
 A quantifier cannot take scope outside of the finite clause in which it is located.

Apparent exceptions to the generalization in (4) have been noted in the literature in which a quantifier seems to be able to take scope outside of the (subjunctive) finite clause in which it appears (call this scope interpretation *extra wide scope*). We present the results of an acceptability rating study conducted in Hungarian whose goal is to empirically investigate whether such extra wide scope readings are indeed systematically available, and if so, whether this is due to the syntactic transparency of subjunctive clauses, a view defended in Wurmbrand (2013, 2018), or it can be ascribed to a process of semantic reanalysis, as claimed in Farkas and Giannakidou (1996).

The structure of the paper is as follows. In section 2 we lay out the conundrum in more detail, and briefly review the alternative analytical options proposed by Wurmbrand (2013, 2018) and by Farkas and Giannakidou (1996). Section 3.1 presents the method used in our experiment. Section 3.2 is devoted to a review of the experimental results, and section 3.3 contains their discussion. Finally, Section 4 concludes and points out residual open issues.

## 2. Background and objectives

According to the classic generative view, the locality restriction in (4) is grammatical in nature. It has been proposed in various forms that it originates from syntactic locality conditions on movement (Hornstein 1995; Johnson 2000; Fox 2000; Cecchetto 2004; Bianchi and Chesi 2010; Takahashi 2011; Abe 2017). If any of these approaches is correct, (4) can be expanded to (5).

- (5) *Syntactic approach to the classic finite clause boundedness constraint on quantifier scope*  
 A quantifier cannot syntactically take scope outside of the finite clause in which it is located due to independent principles governing syntax, in particular, principles of locality of movement.

Some apparent exceptions to (4) have been noted, however: in such cases the scope of a quantifier extends beyond the finite clause in which it is located. The type of cases we consider in the present paper is illustrated by Farkas and Giannakidou (1996) and Wurmbrand (2013, 2018) with examples similar to (6):

- (6) A student made sure that each professor had a ride to his hotel.

On the extra wide, inverse scope reading of the universal quantifier, for each professor there is a (potentially different) student who made sure that the professor had a ride to his hotel. If the latter interpretation is indeed an available reading of (6), then it constitutes an exception to (4).

At the same time, this would beg the question how to account for the apparent difference between (6), involving a subjunctive complement clause, and (3b), which involves an indicative complement clause. One possibility is to maintain the weakened generalization in (7):

- (7) *Weak finite clause boundedness constraint*

Due to principles governing the locality of syntactic movement, a quantifier cannot syntactically take scope outside of the finite clause in which it is located, with the exception of subjunctive complement clauses such as in (6).

Wurmbrand (2013) develops a formal syntactic approach along the lines of (7): one that is more permissive than (5) in allowing for QPs to raise out of (some) subjunctive clauses but still prohibiting QR from an indicative clause.<sup>1</sup> The key idea of the approach, which we cannot review here for reasons of space, is that because subjunctive complement clauses are selected for their mood property, they do not qualify as a local bounding domain (a phase, in terms of Chomsky 2001) for QR. On this approach, apparent exceptions such as (6) are licensed syntactically.

By contrast, Farkas and Giannakidou (1996) propose a semantic approach to extra wide scope in sentences like (6). According to their account extra wide scope interpretation in such examples is not to be accounted for in syntax, but in semantics, specifically, by way of what we will call here semantic reanalysis. Due to this reanalysis, the indefinite existential quantifier, the subject of the matrix clause (NP1), and the universal quantifier, the subject of the embedded clause (NP2), can be reinterpreted as co-arguments of what we may call a complex predicate formed by the matrix and the embedded predicate. This output can be represented schematically as follows:<sup>2</sup>

## (8) NP1 (x made sure y had a ride) NP2

A prerequisite of the proposed semantic reanalysis is that the complex two-place predicate (formed by the matrix and the subordinate predicates) should have the matrix NP1 as the initiator, while the subordinate NP2 as the affected argument in the complex eventuality made up of eventuality e1, denoted by the matrix predicate, and eventuality e2, denoted by the embedded predicate. The semantic relation of extended co-argumenthood then forms the basis of Farkas and Giannakidou's semantic account of inverse scope in sentences like (6) (formulated in terms of Farkas's (1997) indexical theory of scope).

Farkas and Giannakidou's take on examples like (6) can be summarized as in (9):

- (9) *Classic finite clause boundedness constraint + semantic reanalysis*  
 A quantifier cannot normally take scope outside of the finite clause in which it is located. Extra wide scope out of subjunctive clauses like in (6) are derived through semantic reanalysis.

Note that (9) does not implicate any *syntactic* reanalysis (for instance, clausal restructuring). The extra wide scope in (6) is not available by means of long QR, in conformity to (5).

These empirical and analytic proposals motivated our acceptability rating study, to be presented in this paper. While quantifier scope has received a fair amount of attention in empirical work, scope interpretations in sentences similar to (6) have not been adequately treated experimentally.<sup>3</sup> The present experimental study, based on data from Hungarian, contributes to filling this empirical gap.

One objective of the paper is to address question (10a). A second objective, which can be formulated only if (10a) is answered in the affirmative, is to explore scope interpretation in structures that are expected to license extra wide inverse scope on a semantic reanalysis based account (9), but which are predicted to disallow long QR even under the weakened syntactic approach in (7). Such structures are of interest because they have the potential to adjudicate between the syntactic and the semantic approach. We rely on a strong island to create the construction required to test this question. As QR is a movement operation, if extra-wide scope is obtained by QR, then it should not be available out of strong island. The second question can then be formulated as in (10b). If results show that (10b) too is to be answered in the affirmative, then that would support a semantic reanalysis based account (9) over the syntactic account along the lines of

(7). If the semantic account takes care of extra wide scope in sentences like (6), then such examples do not motivate the weakening of the classic syntactic approach in terms of a finite clause bounded QR (= (6)) to (7). Conversely, if our findings indicate a negative answer to (10b), then that would favour the weak syntactic approach in (7) over the semantic approach in (9), since the weak syntactic approach predicts extra wide scope in (6) without licensing extra wide scope out of a syntactic island.<sup>4</sup>

- (10) a. Do judgment data collected from native speakers of Hungarian confirm or disconfirm that the extra wide scope of the universal quantifier is available in sentences like (6)?
- b. Is an extra wide scope reading available in sentences similar to (6) in which the embedded subjunctive clause is in a strong syntactic island, while the conditions for semantic reanalysis are met?

### 3. Experiment

#### 3.1 Method

To address the empirical questions in (10a) and (10b) above, we designed a rating experiment. In the experiment, each experimental item included a context-setting question and an answer. The answer was a complex sentence that contained both a universal QP and a existential QP (namely, a singular noun phrase headed by an indefinite article). The participants' task was to rate a given scope interpretation of the answer on a seven-point Likert scale (ranging from 'inappropriate' to 'appropriate').

Answers in the critical items were constructed in a three-by-two design. The first factor, Scope, had two levels: (i) the universal QP either had Linear scope or (ii) it had Inverse scope over the existential QP. The second factor, called Boundary, had three levels. (i) In Null Boundary condition, the universal QP and the existential QP were located in the same (finite) clause. (ii) In Finite Boundary condition, one of the QPs was in the subjunctive clause that functioned as the complement of the matrix verb, while the other QP was in the matrix clause. (iii) In the Island Boundary condition, the subjunctive subordinate clause was part of a complex NP island. The complex NP was the complement of the matrix verb, and the subordinate clause was the complement of the head noun of the complex NP.

The context-setting questions were of an invariant structure in all conditions. Each condition had six lexicalizations. (11) provides sample stimuli of the Linear Scope condition, while (12) illustrates the Inverse

Scope condition. The universal quantifier was a subject in all six experimental conditions. In the Null Boundary condition (both in its Linear and in its Inverse targeted scope reading), in which the two QPs were co-arguments in the same clause, the existential QP was an object. In the other four conditions the existential QP was the subject of a subjunctive complement clause.

(11) **Linear Scope**

- a. Question: ‘How do the managers make sure that each celebrity should get home before midnight?’
- b. Null B.: A főnök úgy gondolja,  
the chief so thinks  
[hogy a rendezvény után **mindegyik testőr**  
that the event after each bodyguard  
juttasson haza **egy hírességet** időben].  
take.SUBJ home a celebrity.ACC in.time  
‘The chief thinks [that **each bodyguard** should take home  
**a celebrity** in time after the event].’
- c. Finite B.: A rendezvény után **mindegyik testőr** elintézi,  
the event after each bodyguard arranges  
[hogy hazajusson **egy híresség** időben].  
that home.get.SUBJ a celebrity in.time  
‘After the event **each bodyguard** arranges [that **a**  
**celebrity** should get home in time].’
- d. Island B.: A rendezvény után **mindegyik testőr** elintézi  
the event after each bodyguard arranges  
[azt a megbízást,  
that.ACC the task.ACC  
[hogy hazajusson **egy híresség** időben]].  
that home.get.SUBJ a celebrity in.time  
‘After the event **each bodyguard** arranges [the task  
[that **a celebrity** should get home in time]].’

(12) **Inverse Scope**

- a. Question: ‘How do the managers make sure that each celebrity should get home before midnight?’

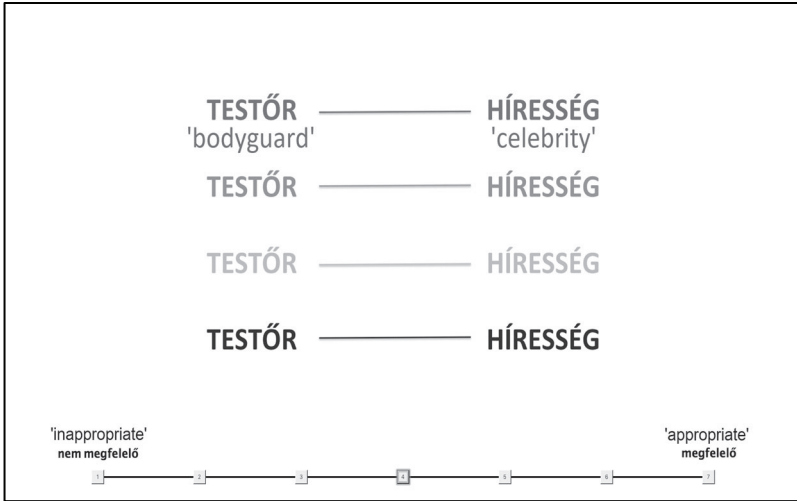
- b. Null B.: A főnök úgy gondolja,  
 the chief so thinks  
 [hogy a rendezvény után **egy testőrt**  
 that the event after a bodyguard.ACC  
 béreljen fel **mindegyik híresség** a hazajutáshoz].  
 hire.SUBJ up each celebrity the getting.home.for  
 ‘The chief thinks [that **each celebrity** should hire **a bodyguard** to get home in time after the event.]’
- c. Finite B.: A rendezvény után **egy testőr** intézi el,  
 the event after a bodyguard arranges away  
 [hogy hazajusson **mindegyik híresség** időben].  
 that home.get.SUBJ each celebrity in.time  
 ‘After the event **a bodyguard** arranges [that **every celebrity** should get home in time].’
- d. Island B.: A rendezvény után **egy testőr** intézi el  
 the event after a bodyguard arranges away  
 [azt a megbízást,  
 that.ACC the task.ACC  
 [hogy hazajusson **mindegyik híresség** időben]].  
 that home.get.SUBJ each celebrity in.time  
 ‘After the event **a bodyguard** arranges [the task [that **each celebrity** should get home in time]].’

All question–answer dialogues were pre-recorded with natural intonation in a soundproof studio. The targeted scope reading (Linear or Inverse) was displayed in a schematic diagram, which appeared on a computer screen. The elements quantified over by the quantifier that had wider scope consistently appeared on the left hand side of each diagram, while the elements quantified over by the dependent, narrow scope quantifier were arranged on the right hand side. The distributive mapping between the two sets of elements was represented by connecting lines. Each line, along with the left-hand side and right-hand side element that it connected, appeared in a different colour. This colour-coding (to which we explicitly drew participants’ attention in the task instructions) served to emphasize that each pair of elements, and indeed each individual element, is distinct.

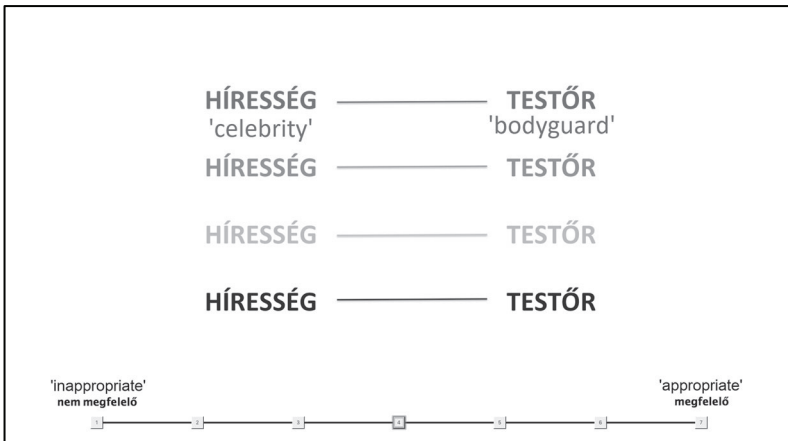
As mentioned above, on the targeted interpretation of the examples in (11) the universal quantifier (quantifying over bodyguards) had Linear Scope over the existential quantifier (quantifying over celebrities). Figure 1



shows the visual stimulus that represented this interpretation. Figure 2, on the other hand, contains the display that represented the Inverse Scope interpretation of examples in (12).



**Fig. 1: Linear Scope visual stimulus accompanying (11)**



**Fig. 2: Inverse Scope visual stimulus accompanying (12)**

The experiment contained a Control condition. Items in this condition were virtually identical to the Inverse Scope Island Boundary condition, except that in this condition the universal QP was an object. On Farkas and Giannakidou’s (1996) semantic approach to extra wide scope in sentences like (6), only the matrix and the embedded subject can ever become extended co-arguments (because only the embedded subject can function as an affected extended co-argument; see Section 1), therefore we expected these items to be rated low on an inverse scope interpretation. (12) presents a sample of the Control condition, while Figure 3 provides the corresponding diagram.

(13) a. Question: ‘How are the planes checked during the spring testing period at the airport?’

b. Island B.: A gyakorlat után **egy pilóta** adja ki  
 the test after a pilot gives out  
 [azt a feladatot,  
 that.ACC the task.ACC  
 [hogy a karbantartók nézzék át  
 that the repairmen check.SUBJ through  
**mindegyik gépet** a biztonság kedvéért]].  
 each plane.ACC the safety sake.for  
 ‘After the testflight **a pilot** orders [the task [that the repairmen should check **each plane** for the sake of safety]].’

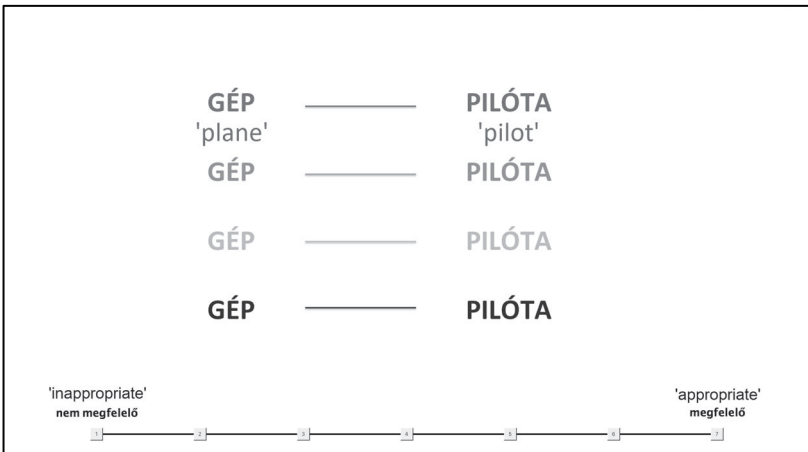


Fig. 3: Sample Control stimulus (13)

Fifteen additional filler dialogues were included to balance the anticipated pattern of judgments across the experiment. Taken together with the six lexicalizations of the six critical conditions, and the three items of the Control condition, the total number of test items was 54.

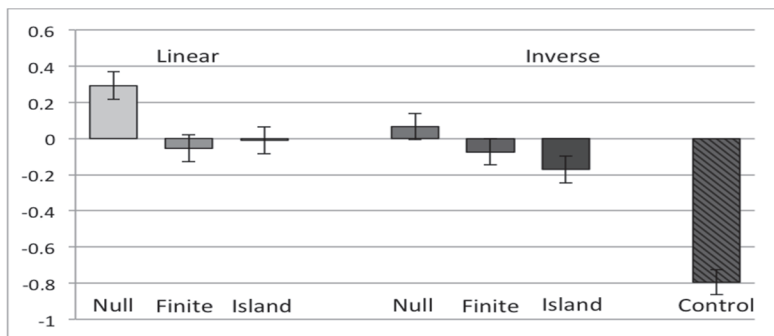
The experimental procedure was as follows. Trials were presented in one of two pseudo-randomized orders to each participant. In each trial, the audio stimulus was first played without any visual stimulus. Then the diagram appeared on the screen, followed by a five-second pause. After this the recording was played another time, with the diagram still displayed. As soon as the second repetition of the audio was over, a seven-point Likert scale appeared at the bottom of the screen. The diagram and the scale then remained on the screen until the judgment was entered.

The experiment was scripted in Inquisit Software, and it was carried out online. In order to familiarize the participants with the task, the test trials were preceded by a short practice session containing items unrelated to the experimental items.

Fifty-eight adult native speakers participated in the experiment; two of them were excluded because they reported that their parents were not native speakers of Hungarian. Of the remaining fifty-six participants (mean age: 36 years), 43 were female and 13 were male. The ratio of the participants who filled in the trials in the two lists was 22:34.

### 3.2. Results

Figure 4 depicts the mean values of the  $z$ -transformed ratings (transformed by subject).<sup>5</sup> As shown, the Control condition received very low values. On the other hand, the critical conditions were judged around the middle of the scale.



**Fig. 4: Mean values of  $z$ -transformed judgments (+/- Standard Error)**

The Linear Scope Null Boundary condition got the highest mean score ( $N=168$ ,  $M=0.291$ ,  $SE=\pm 0.08$ ); while the Control condition had the lowest  $z$ -score mean ( $N=168$ ,  $M=-0.794$ ,  $SE=\pm 0.07$ ). The mean of the other conditions was around the overall mean judgment: Linear Scope Finite Boundary ( $M(168)=-0.055$ ;  $SE=\pm 0.07$ ); Linear Scope Island Boundary ( $M(168)=-0.009$ ;  $SE=\pm 0.08$ ); Inverse Scope Null Boundary ( $M(168)=0.066$ ;  $SE=\pm 0.07$ ); Inverse Scope Finite Boundary ( $M(168)=-0.073$ ;  $SE=\pm 0.07$ ); Inverse Scope Island Boundary ( $M(168)=-0.171$ ;  $SE=\pm 0.08$ ).

The statistical analysis of the  $z$ -transformed dataset (including the Control condition) was conducted in R with the `lmer` function, based on linear mixed effects models.<sup>6</sup> After model selection, using a backward elimination method starting with the full model, the final model consisted of the two fixed factors with interaction between them, and the item and the subject as random factors (the latter had the two fixed factors as random slopes, with interaction). This model found that the Boundary Factor had a significant main effect ( $X^2(3)=35.78$ ;  $p<0.001$ ), while the Scope Factor did not ( $X^2(1)=2.19$ ;  $p=0.14$ ). No interaction was found between the two main factors ( $X^2(2)=1.08$ ;  $p=0.58$ ).

Pairwise *post hoc* (Tukey) comparisons of the conditions revealed no significant difference either between the Linear Scope Null Boundary and Finite Boundary conditions, or between the former and the Linear Scope Island Boundary condition. No significant difference was detected between the Linear Scope Finite Boundary and Island Boundary conditions either. Similarly, the Inverse Scope conditions did not differ from each other pairwise. On the other hand, each Inverse Scope condition differed significantly from the Control condition with a considerable effect size (Cohen's  $d$ ): Null Boundary vs Control ( $t$ -ratio=5.31;  $p<0.001$ ;  $d(\text{Cohen})=0.94$ ); Finite Boundary vs Control ( $t$ -ratio=4.31;  $p<0.001$ ;  $d(\text{Cohen})=0.79$ ); Island Boundary vs Control ( $t$ -ratio=3.79;  $p<0.01$ ;  $d(\text{Cohen})=0.67$ ). Lastly, there was no significant difference between the Linear Scope and Inverse Scope at any of the three levels of the Boundary factor.

### 3.3. Discussion

The experiment was designed to address the questions formulated in (10a) and (10b), repeated here:

- (10) a. Do judgment data collected from native speakers of Hungarian confirm or disconfirm that the extra wide scope of the universal quantifier is available in sentences like (6)?
- b. Is an extra wide scope reading available in sentences similar to (6) in which the embedded subjunctive clause is in a strong syntactic island, while the conditions for semantic reanalysis are met?

In our experiment the item corresponding to sentence type (6) contained an extra wide scope taking universal quantifier embedded in a subjunctive clause. The relevant item, representing the Inverse Finite condition, is repeated here for convenience:

- (12) c. A rendezvény után **egy testőr** intézi el,  
 the event after a bodyguard arranges away  
 [hogy hazajusson **mindegyik híresség** időben].  
 that home.get.SUBJ each celebrity in.time  
 ‘After the event **a bodyguard** arranges [that **every celebrity** should get home in time].’

To begin with the first question (10a), the results of our experiment answer confirm the availability of extra wide scope in sentences similar to (6). This is due to the fact that no significant difference was found between the linear and inverse scope interpretations of the Finite Boundary condition. The Linear Scope Finite Boundary condition is predicted by all extant grammatical theories of quantifier scope to be licensed. Since the acceptability of this condition is not significantly different from that of the inverse scope reading in the Finite Boundary condition (the condition corresponding to (6)), we conclude that the latter scope reading is also licensed. This conclusion is reinforced by two further findings. First, the Inverse Scope Finite Boundary condition also did not differ from the Inverse Scope Null Boundary condition, in which inverse scope was clause-bound and was therefore grammatically licensed. Second, inverse scope in the Control condition, which was expected to be ungrammatical both on Farkas and Giannakidou’s (1996) and on Wurmbrand’s (2013) theory, was significantly worse than the Inverse Scope Finite Boundary condition, with a large effect size ( $d = 0.79$ ).

With regard to the second research question (10b), concerning extra wide scope in constructions in which the dependent subjunctive clause is part of a complex NP (=Inverse Scope Island Boundary condition), our results suggest that it is available just as it is in sentences like (6) (=Inverse

Scope Finite Boundary condition), with no significant difference between ratings. Once again, inverse scope in the Control condition, which was expected to be ungrammatical both on Farkas and Giannakidou's (1996) and on Wurmbrand's (2013) theory, was significantly worse than the Inverse Scope Finite Boundary condition, with a considerable effect size ( $d = 0.67$ ). That the Control condition can be taken to fall into the category of ungrammatical scope interpretations is supported by Sprouse et al.'s (2017) conjecture that syntactic island effects, which lead to ungrammaticality, yield 0.5 standard deviations below the middle of the range of acceptability. The fact that the Control condition produced a z-score of -0.79 can be taken to indicate that the targeted scope reading in that condition is indeed ungrammatical, as opposed to the one in the Inverse Scope Island Boundary condition, which yielded a z-score of -0.17, and which can then be taken to be grammatically licensed. The fact that inverse scope is licensed out of a complex NP island in cases in which the conditions for semantic reanalysis are met are accurately predicted by Farkas and Giannakidou's (1996) approach. By contrast, the difference between the Inverse Scope Island Boundary condition and the Control condition is unexpected on Wurmbrand's (2013) weak syntactic account, on which Quantifier Raising out of a complex NP island is plainly disallowed.

Finally, our results have confirmed the difference between the subject and the object of the embedded subjunctive clause in their ability to license extra-wide scope when the subjunctive clause appeared in a complex NP island: the Inverse Scope Island Boundary condition was significantly more acceptable than the Control condition. This is of interest because Hungarian lacks some of the structural subject-object asymmetries that are well known configurational languages (É. Kiss 2002). This finding, then, provides indirect confirmation for Farkas and Giannakidou's (1996) assumption that this subject-object asymmetry is not due to a syntactic asymmetry, but rather, it derives from a semantic difference in terms of extended co-argumenthood: the embedded subject functions an extended co-argument of the matrix subject (while the embedded object does not), because it functions as an affected participant in the complex event denoted by the matrix and embedded predicates together.<sup>7</sup>

#### 4. Conclusions and open questions

This paper presented a rating experiment which investigated the status of extra wide scope out of subjunctive complement clauses, which have been recurrently cited as evidence against the strict and uniform locality of Quantifier Raising (i.e., its finite clause-boundedness). The study substantiates

two conclusions. First, based on judgments of sentences like (6) in Hungarian, it empirically confirms that wide scope out of a subjunctive complement clause is indeed available. This testifies that the purported phenomenon of inverse quantifier scope across a subjunctive clause boundary is real, and it extends beyond the Indo-European language family. This is in line with the grammatical accounts of Farkas and Giannakidou (1996) and Wurmbrand (2013).

Second, our results reveal that the same extra wide inverse scope reading is also available in sentences minimally different from (6), in which the dependent subjunctive clause forms part of a complex NP island. Farkas and Giannakidou's (1996) account in terms of what we have referred to as semantic reanalysis captures the availability of extra wide inverse scope out of complex NP islands in the case at issue (in which the conditions of semantic reanalysis are met), while Wurmbrand's (2013) weak syntactic approach does not. Because the same semantic reanalysis based account also covers extra-wide inverse scope out of simple subjunctive complement clauses in sentences like (6), we conclude that this sentence type does not warrant the weakening of the syntactic approach to the classic finite clause boundedness constraint on quantifier scope in (4) proposed in Wurmbrand (2013). In other words, examples like (6) do not compel us to abandon a uniform syntactic definition of what counts as a local domain for Quantifier Raising (and also for unbounded A-bar movements), according to which for these purposes indicative and subjunctive clauses can be treated symmetrically.

The objectives of the present contribution have been modest: they are limited to making these two main points. A number of related issues have been left open here. First and foremost, our conclusions regarding the basic locality of Quantifier Raising do not mean to imply that QR cannot exceptionally be longer than finite clause bounded in certain well-defined cases. These may include Antecedent Contained Deletion constructions (Cecchetto 2004, Syrett 2015), and object QR out of a finite complement clause that has a bound pronoun as a subject (Grano and Lasnik's 2018 Bound Pronoun Effect).

Second, we also leave open here the precise mechanism by which semantic reanalysis is made available. Farkas and Giannakidou (1996) proposed to adopt Farkas's (1996) indexical theory of scope. This theory is technically not incompatible with the classic model in (4) that incorporates Quantifier Raising. Nevertheless, any alternative theory that is able to capture Farkas and Giannakidou's notion of 'extended co-argumenthood' and can implement scopal commutativity of co-arguments would also be serviceable.

## References

- Abe, Jun. 2017. *Minimalist Syntax for Quantifier Raising, Topicalization and Focus Movement: A Search and Float approach for Internal Merge*. Dordrecht: Springer.
- Beghelli, Filippo. 1993. "A minimalist approach to quantifier scope". *Proceedings of the North East Linguistics Society 23*, edited by Amy J. Schafer, 65–80. Amherst: GLSA.
- Bianchi, Valentina, and Cristiano Chesi. 2010. "Reversing the perspective on Quantifier Raising". *Rivista di Grammatica Generativa 35*: 3–38.
- Cecchetto, Carlo. 2004. "Explaining the locality conditions of QR: Consequences for the theory of phases". *Natural Language Semantics 12* (4): 45–397.
- Chomsky, Noam. 1976. "Conditions on Rules of Grammar". *Linguistic Analysis 2* (4): 303–351.
- Chomsky, Noam. 2001. "Derivation by Phase". In *Ken Hale: A Life in Language*, edited by Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.
- É. Kiss, Katalin. 2002. *Syntax of Hungarian*. Cambridge: Cambridge University Press.
- Farkas, Donka. 1981. "Quantifier scope and syntactic islands". In *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society*, edited by Roberta A. Hendrick, Carrie S. Masek, and Mary Frances Miller, 59–66. Chicago: University of Chicago, Chicago Linguistic Society.
- Farkas, Donka. 1996. "Evaluation Indices and Scope". In *Ways of Scope Taking. Studies in Linguistics and Philosophy 65*, edited by Anna Szabolcsi, 183–215. Dordrecht: Springer.
- Farkas, Donka, and Anastasia Giannakidou. 1996. "How clause-bounded is the scope of universals?". In *Proceedings of SALT VI*, edited by Teresa Galloway and Justin Spence, 35–52. Ithaca: CLC Publications, Cornell University.
- Fodor, Janet D., and Ivan Sag. 1982. "Referential and quantificational indefinites". *Linguistics and Philosophy 5*: 355–398.
- Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.
- Fox, Danny and Uli Sauerland. 1996. "Illusive scope of universal quantifiers". In *Proceedings of the North East Linguistics Society 26*, edited by Kyiomi Kusumoto, 71–85. Amherst: University of Massachusetts, GLSA.



- Grano, Thomas, and Howard Lasnik. 2018. "How to neutralize a finite clause boundary: Phase theory and the grammar of bound pronouns". *Linguistic Inquiry* 49 (3): 465–499.
- Hornstein, Norbert. 1995. *Logical form: From GB to Minimalism*. Oxford: Blackwell.
- Johnson, Kyle 2000. "How far quantifiers go?". In *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, edited by Roger Martin, David Michaels, and Juan Uriagereka, 187-210. Cambridge: The MIT Press.
- May, Robert. 1977. *The grammar of quantification*. PhD diss., Massachusetts Institute of Technology.
- May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, MA: MIT Press.
- Ross, John Robert 1967. "Constraints on Variables in Syntax." PhD diss., MIT Cambridge, Massachusetts.
- Sprouse, Jon, and Diogo Almeida. 2017. "Design sensitivity and statistical power in acceptability judgment experiments". *Glossa* 2 (1): 1–32.
- Syrett, Kristen. 2015. "Experimental support for inverse scope readings of finite-clause embedded Antecedent-Contained Deletion sentences". *Linguistic Inquiry*, 46 (3): 579–592.
- Szabolcsi, Anna 1997. "Strategies for scope taking". In *Ways of Scope Taking*, edited by Anna Szabolcsi, 109-154. Dordrecht: Kluwer Academic Press.
- Takahashi, Masahiko. 2011. "Some theoretical consequences of Case-marking in Japanese." PhD diss., University of Connecticut, Storrs, CT.
- Tanaka, Misako. 2015a. "Asymmetries in long-distance QR". In *Proceedings of Forty-First Annual Meeting of the Berkeley Linguistics Society*. Vol. 41, edited by Anna E. Jurgensen, Hannah Sande, Spencer Lamoureux, Kenny Baclawski, and Alison Zerbe, 493-501. Berkeley: Berkeley Linguistics Society.
- Tanaka, Misako. 2015b. *Scoping out of Adjuncts: Evidence for the Parallelism between QR and Wh-movement*. PhD diss., University College London.
- Wurmbrand, Susi. 2013. "QR and selection: Covert evidence for phasehood". In *Proceedings of the North East Linguistics Society 42*, edited by Stefan Keine and Shayne Sloggett, 619–632. Amherst: University of Massachusetts, GLSA
- Wurmbrand, Susi. 2018. "The cost of raising quantifiers". *Glossa* 3(1):19: 1–40.

## Notes

---

<sup>1</sup> Wurmbrand (2018) accounts for the same difference by proposing a processing-based explanation that relies on structural differences assumed between subjunctives and indicatives.

<sup>2</sup> We diverge from Farkas and Giannakidou's (1996) formulations here. The scheme in (8) is our own. Both (8) and the notion of 'complex predicate' are used for presentational purposes. Further, Farkas and Giannakidou do not call the interpretational option by which NP1 and NP2 are understood as co-arguments in some complex eventuality 'semantic reanalysis'; they simply take the semantic relationship between NP1 and NP2 to be licensed by virtue of the semantics of the matrix and embedded predicates. Readers should consult Farkas and Giannakidou (1996: 38-42) for the original formulations.

<sup>3</sup> Tanaka (2015a; 2015b) is an important exception. For reasons of space, we cannot discuss her methods and results here. We merely note here that Tanaka's findings confirm the difference between subjunctive and indicative complement clauses that is assumed both in Farkas and Giannakidou (1996) and in Wurmbrand (2013, 2018) in that subjunctives allow extra wide scope across their boundary significantly more readily than do indicative clauses.

<sup>4</sup> In difference to the semantic and syntactic accounts mentioned in this review, a third type of approach that has been suggested of extra wide quantifier scope posits that quantifier scope is in fact unbounded (like other A-bar movements). Any restrictions that may be found are extra-grammatical in nature: they arise due to processing demands (for an insightful recent implementation, see Wurmbrand 2018). In the present paper we do not comment on this general approach, apart from noting that it is not incompatible with the main results of the experiment we present in the next section.

<sup>5</sup> Z-scores are standardized scores corresponding to the number of standard deviations that a given raw score is above or below the mean (which is represented by  $z=0$ ). We estimated means and standard deviations for each subject based on the responses across all test items.

<sup>6</sup> Mixed models take into account both fixed effects (experimental manipulations) and random effects (the effect of cross-item and cross-subject differences), and their interactions, within the same analysis.

<sup>7</sup> The finding is also of interest because Tanaka (2015a, 2015b) failed to detect a significant difference in English between the availability of scope-taking by subject and object QPs out of subjunctive as well as indicative complement clauses.

# CHAPTER SEVEN

## POLARITY SENSITIVE ITEMS IN NON-VERIDICAL CONTEXTS IN ROMANIAN

MIHAELA ZAMFIRESCU

The aim of the present paper is to investigate non-veridical contexts in Romanian and examine the behaviour of polarity-sensitive items *deja* ('already'), *încă* ('still') and *doar* ('only') in such contexts. A subsequent aim of this investigation is to see if we can establish a hierarchy of polarity sensitive items in Romanian based on the criterion of compatibility of the abovementioned items with types of negative strength.

Keywords: PSI (polarity sensitive item) *any*, hierarchy, negative strength, positive polarity

### 1. Introduction and background data

The present paper aims at investigating non-veridical contexts in Romanian and testing whether items that exhibit sensitivity to positive polarity, like *deja* ('already'), *încă* ('still') and *doar* ('only') are allowed in such contexts since one of the simplest definitions one can propose for defining items sensitive to positive polarity would be, in broad terms, that: positive polarity items are items whose occurrence is legitimate only in assertive contexts, as shown in (1a) where we exemplify the case of *something* in English. Consequently, positive polarity items' occurrence in negative contexts should not be legitimate, as shown in (1b). Unfortunately, things do not stand as such at all times, as we shall see in what follows. Not only can (1b) be interpreted as correct if we consider it a case of emphatic denial<sup>1</sup>, as in (1c), but the sentence is correct if *some* is interpreted to scope above negation, as in (1d). Also, PPIs are legitimate scoping below superordinate negation, as in (2a) and the relation between negation and PPI can be subject to intervention effects and thus we see that PPIs are legitimate

under negation if operators such as *always* intervene between negation and the PPI, as in (2b).

- (1) a. Sylvia has *something* to write as homework for the English class.  
 b. \*Sylvia doesn't have *something* to write as homework for the English class.  
 c. Sylvia has *something* to write as homework for the English class.  
 Wrong! Sylvia DOES NOT have *something* to write as homework for the English class.  
 d. Sylvia has *something* to write as homework for the English class. She doesn't want to write *some* of the exercises.
- (2) a. I don't think that Sylvia has *something* to write as homework for the English class.  
 b. She doesn't always have *something* to write as homework for the English class.

Another aim of this paper is to see if we can propose a hierarchy for the grammatical positive polarity items under investigation, starting from the criterion of their occurrence or impossibility of occurrence in certain negative contexts.

## 2. Licensing mechanisms

In the vast literature on negative polarity, one of the questions that has been at the core of most research studies is: what is the mechanism that can explain how polarity sensitive items are licensed? A number of proposals have been put forward in order to answer the previously mentioned question, some of the most notable invoking arguments of syntactic, semantic or pragmatic nature. Syntactic accounts, starting with Klima (1964), have focused on the relation between negation and the negative polarity item (NPI) and proposed that in order for a negative polarity item to be licensed it needs to be in the immediate scope of negation, which, in broad terms could be understood as a c-command relation between negation, the licenser and the negative polarity item, the licensee. In the following examples, *one bit*, whose presence is legitimate only in the scope of negation, is felicitous in example (3a) as it is licensed in the immediate scope of negation and is conversely infelicitous in example (3b) because negation does not license NPIs across the boundary of a clause. *One bit*, in

English, is the type of negative polarity item whose occurrence is infelicitous unless it is in the narrow domain of negation.

- (3) a. I didn't like her one bit.  
 b. \*It's not customary that I fall over backwards one bit to help them.
- (4) a. Simona hardly said anything to me after our quarrel.  
 b. Implicature: Simona almost didn't say anything to me after our quarrel.

Since syntactic accounts could not fully explain by means of the c-command relation how negative polarity items were licensed in cases where negation was not overt without postulating licensing by means of pragmatic relations, such as the negative implicatures generated by the items themselves, as Linebarger (1987) proposed for examples such as the one under (4), semanticists proposed explanations based on the semantic property of downward-entailment. This semantic property of downward-entailment is the property of contexts to license inferences from sets to subsets. Ladusaw (1979: 59-61) proposed a reinterpretation of the c-command relation that syntacticians employed, postulating that an expression *X* is in the scope of another relation *Y*, which basically is still the c-commanding relation, if any expression contained in *X* is also in the domain of *Y*. At the foundation of the property of downward-entailment lies the notion of monotonicity, where, as defined in Szabolcsi (2010: 54) "a function *f* is monotonically decreasing with respect to a particular argument iff it reverses the partial ordering in its domain". In other words, "if *X*, *Y*, are in the domain of *f* and  $X \leq Y$ , then  $f(Y) \leq f(X)$ " (Szabolcsi (2010: 54)). 'Not' is downward monotonic and not upward monotonic, as the following example under (5a) demonstrates, by contrast with the example under (5b):

- (5) a. Simona didn't meet anyone famous at the party. → Simona didn't meet Madonna at the party.  
 b. Simona didn't meet Madonna at the party. /→ Simona didn't meet anyone famous at the party.

Pursuing this line of argumentation, that negation bears the property of downward-entailment, which can license inferences from the set *sports* to the subset *skiing*, the negative polarity item *any* is correctly licensed in the following example.

- (6) Sylvia didn't practice any sports this weekend. → Sylvia didn't practice any skiing.

Unfortunately, despite the ingenuity of Ladusaw's (1979) proposal, there were still examples that could not be accounted for by the downward-entailment theory. One such example is the operator *most*. As demonstrated by the following example, *most* is not downward entailing, since whenever the sentence under (7) is true by replacing the predicate in (7) with a more limited predicate, a subset of the initial predicate the resulting sentence is not necessarily true. It does not necessarily follow that the children who have eaten chocolate and loved it will also enjoy mint chocolate, where mint chocolate is a subset of the class chocolate, since it is more limited, exclusive.

- (7) Most children who have eaten chocolate love it. /→ Most children who have eaten mint chocolate love it.

The following example demonstrates that *most* is neither upward-entailing, since whenever the sentence under (8) is true, by replacing the predicate in (8) with a broader predicate, a superset of the initial predicate the resulting sentence is not necessarily true. It does not necessarily follow that if raptors fly well, so will all birds, even if *birds* is the superset of *raptors*. It is common knowledge that not all birds can fly.

- (8) Most birds of prey fly well. /→ Most birds fly well.

Consequently, *most* can only be described as a non-monotonic operator. Nevertheless, it does license NPIs, as shown in the next example.

- (9) Most tenants with the least bit of decency don't drill late at night.

Trying to see what semantic property contexts that license polarity sensitive items share, Zwarts (1995) proposed a hierarchy based on the criterion of negative strength, meaning which of those contexts evinced a stronger degree of negativity and thus are more restrictive in allowing or disallowing polarity sensitive items in their domain. The most restrictive type of context, allowing the smallest number of polarity sensitive items in their domain, is the context generated by *not*, which is an anti-morphic operator. An operator Op is anti-morphic<sup>2</sup> if and only if Op(A) and Op(B) is equivalent to Op(A or B) and Op(A) or Op(B) is equivalent to Op(A and B). For example, *Simona did not sell old magazines and Simona did not buy*

*old magazines* is equivalent to *Simona did not sell or buy old magazines*; and *Simona did not sell old magazines or Simona did not buy old magazines* is equivalent to *Simona did not (both) sell and buy old magazines* as in (10 a, b). Similarly, *nu* (not), in Romanian, is an anti-morphic operator, in (10 c, d).

- (10) a. Simona did not sell old magazines and Simona did not buy old magazines.  $\leftrightarrow$  Simona did not sell or buy old magazines.
- b. Simona did not sell old magazines or Sylvia did not buy old magazines.  $\leftrightarrow$  Simona did not (both) sell and buy old magazines.
- c. Simona nu gătește pește și Simona nu gătește tocane.  
Simona not cooks fish and Simona not cooks stews.  
“Simona does not cook fish and Simona does not cook stews.”  
 $\leftrightarrow$   
Simona nu gătește pește sau tocane.  
Simona not cooks fish or stews.  
“Simona does not cook fish or stews.”
- d. Simona nu gătește pește sau Simona nu gătește tocane.  
Simona not cooks fish or Simona not cooks stews.  
“Simona does not cook fish or Simona does not cook stews.”  
 $\leftrightarrow$   
Simona nu gătește nici pește nici tocane.  
Simona not cooks neither fish nor stews.  
“Simona does not cook fish and stews.”

Keeping in mind the criterion of how strong the force of negation is, a less restrictive class of negative contexts is the class of anti-additive operators. Typical examples of anti-additive operators are *refuse* and *without*. Such operators allow in their domains not only the polarity sensitive items that are allowed in the scope of anti-morphic operators but also other several types of polarity sensitive items which are not legitimate in the scope of the abovementioned class, that of anti-morphic operators. A function  $f$  is anti-additive if for all  $x, y$  such that  $f(x \cup y) = f(x) \cap f(y)$ . In other words, *Simona refuses to pay taxes or the mortgage* is equivalent to *Simona refuses to pay taxes and Simona refuses to pay the mortgage*<sup>3</sup>. Imagine a woman sued for having incurred a huge amount of debt and what the prosecutor would say:

- (11) Simona refuses to pay taxes or the mortgage. ↔ Simona refuses to pay taxes and Simona refuses to pay the mortgage.

In a similar fashion, *refuză*, in Romanian qualifies as an anti-additive operator, as demonstrated by the following example.

- (12) Simona refuză supa sau legumele crude.  
 Simona refuses soup.DEF or vegetables.DEF raw.  
 Simona refuză supa și Simona refuză legumele crude.  
 Simona refuses soup.DEF and Simona refuses vegetables.DEF raw.  
 Simona refuses the soup or the raw vegetables. ↔ Simona refuses the soup and Simona refuses the raw vegetables.

Since we are talking about a hierarchy of negative strength, the least restrictive type of operator discussed by Zwarts (1995) is the class of downward-entailing operators, downward-entailment being understood in the way Ladusaw (1979) initially described it. Polarity sensitive items that are legitimate in the scope of a downward-entailing operator are also legitimate in the scope of an anti-additive operator and in the scope of an anti-morphic operator, but not the other way around. *Few* and *at most n* are typical examples of downward-entailing operators. In the following examples if the sentence is true then a sentence with a more limited predicate, a subset of the initial predicate is also true. If *few children help with the housework* is true then *few children help with vacuuming* is also true as *vacuuming* is a subset of *housework*.

- (13) a. Sylvia didn't eat any sweets this week. → Sylvia didn't eat any chocolate this week.  
 b. Few children help with the housework. → Few children help with vacuuming.

In the same manner, *puțini* (few) and *cel mult n* (at most n) qualify as downward-entailing operators in Romanian.

- (14) Puțini copiii mănâncă legume. → Puțini copiii mănâncă broccoli.  
 Few children eat vegetables. → Few children eat broccoli.

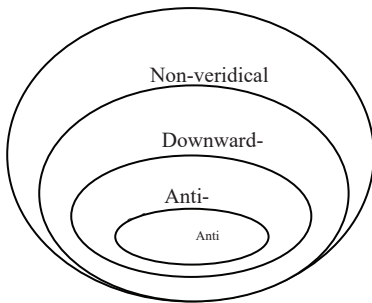
Nevertheless, there are a number of other contexts where the occurrence of negative polarity items is legitimate and we can't account for the grammaticality of these examples neither by invoking a c-command relation as there is no overt negation in such contexts nor by describing them as



downward-entailing, anti-additive or anti-morphic. Comparatives, questions and conditionals are contexts that license polarity sensitive items, as shown in the following examples.

- (15) a. She left earlier than *any* of us had expected.  
 b. Do you take in *any* homeless?  
 c. Please give me a call if there are *any* complaints.

One important semantic study on polarity sensitive items, which gives an account of the previously mentioned examples, was proposed by Giannakidou (1998). Besides the three abovementioned categories, Giannakidou (1998) proposes that there is another fourth category in the hierarchy of negative contexts, whose operators are described as non-veridical operators. Figure 1, borrowed from Hoeksema (2012: 4) perfectly illustrates what Zwarts (1995) meant when saying that the conditions are downwards applicable. As can be seen from this figure, non-veridical contexts are the least strong in terms of the strength of negation. Thus, if a certain negative polarity item is allowed in non-veridical contexts, it is a very weak NPI which is automatically allowed in the rest of the contexts. By contrast, if a negative polarity item is allowed only in anti-morphic contexts, it is a super strong NPI whose occurrence is banned in the rest of the negative contexts.



**Fig. 1: An extended version of Zwarts' (1995) hierarchy of negative strength (Hoeksema 2012: 4)**

The notion of veridicality, as first proposed by Montague (1969), relies on the possibility of obtaining truth-entailments. An operator *F* is veridical if and only if from the truth of *Fp* one can infer that *p* is also true. In case one cannot make such an inference then the operator is non-veridical. For example, in (16a), the necessity operator is veridical since whenever *Fp* is

also true, *p* is also true. By contrast, in (16b), the possibility operator is non-veridical since whenever *Fp* is also true, *p* may or may not be true.

- (16) a. It is necessarily true that Bucharest is the capital of Romania.  
 b. It is possibly true that this pill cures cancer.

As can be seen from the previous example, the truth of propositions, generated by the use of non-veridical operators, is neither asserted nor entailed, as is the case with: questions, imperatives and conditionals.<sup>4</sup>

- (17) a. *Does she love any of them?* does not entail *She loves (at least) one of them.*  
 b. *If she loves him ...* does not entail *she loves him.*

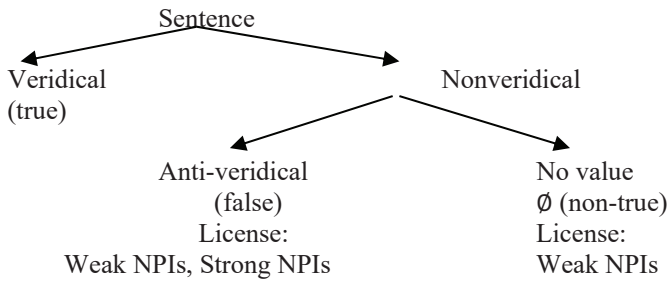
Assertive contexts, for example, are *veridical* because we can claim that they *entail the truth of the proposition*, or in other words, the speaker commits himself/ herself to the truth of the proposition expressed by the sentence. Syntactically negative contexts are, quite the opposite, *anti-veridical* which means that they entail the falsity of the proposition *p*. Truth be told, as was obvious in the previous examples under (16b) and (17), there are sentences which neither entail of the truth of the proposition nor its falsity, which express that the proposition is not necessarily true and which bear the name of *non-veridical*. The following is Giannakidou and Cheng's (2006: 589) proposal for (non)veridicality.

“a. A propositional operator *F* is veridical iff *Fp* entails or presupposes that *p* is true in some individual's epistemic model *ME(x)*; otherwise, *F* is nonveridical.

b. A nonveridical operator *F* is ANTIveridical iff *Fp* entails that NOT *p* in some individual's epistemic model: *Fp* N ! *p* in some *ME(x)*”<sup>5</sup>.

What is essential for studies on polarity and veridicality is that Giannakidou (1998) proposed to describe veridicality relative to the epistemic model of individuals, and thus we are able to apply the concept not only to propositional operators, but to propositions as a whole. In other words, a grammatical affirmative sentence is veridical since its proposition is true relative to the speaker's belief state.

Veridicality can be said to distinguish between what is true, veridical, and non-true, non-veridical, given the speaker's commitment to the truth of the sentences. If we were to propose a diagram, it would look like:



**Fig. 2: Veridicality schema**

One important aspect that need not be forgotten is that any element that is licensed by non-veridicality is also licensed by anti-veridicality but not the other way around. As typical non-veridical contexts, Giannakidou (1998) enumerates the following:

- (18) a. The conditional complementizer *if*, more generally hypothetical clauses.  
If *anybody* should ask for an appointment later than 6 p.m., let me know as soon as possible.
- b. The question complementizer (the Q/wh morpheme), introducing direct or indirect questions.  
Is there *anyone* for Dr. Robbins?
- c. psychological-emotive verbs such as *surprise*, *amaze*, etc.; directive-attitude verbs: want, insist, suggest, allow)  
She's surprised that *anyone* should remember how to light a fire.  
He's amazed that he can read *any* word in French.
- d. Quantifiers: *few*, *little*, *every*, only DP, quantifying adverbials: *whenever*, *wherever*  
Few children have *any* ounce of patience.  
Only Sylvia did *any* of the work.
- e. incorporated negatives: *doubt*, *dislike*, *unlikely*, etc.  
I doubt that he recognizes *anyone*.

- f. negative frequency adverbs: *seldom, rarely*.  
We have rarely heard of *anyone* like him.
- g. *comparatives/superlatives, etc.*  
It's the smartest dog that *anybody* knows.
- h. Relative clauses headed by indefinite or negative determiners like *any, no, every, a*.  
No child who has *any* knowledge of table manners would devour the steak in that way.

### 3. Romanian PPIs in non-veridical contexts

The aim of this section is to investigate whether *deja* ('already'), *încă* ('still') and *doar* ('only') are legitimate in non-veridical contexts such as the ones enumerated in the previous section. Examples under (19), (20) and (21) demonstrate that the previously mentioned positive polarity items are felicitously licensed in non-veridical contexts.

- (19) a. The conditional complementizer *if*, more generally hypothetical clauses.  
Dacă suferiți deja de diabet, atunci este  
If suffer.2PL already of diabetes then is  
important ca împreună cu medicul dumneavoastră să  
important that together with doctor.DEF your SUBJ  
mențineți nivelul glucozei din sânge la  
maintain level.DEF glucose.DEF.GEN from blood at  
cât mai apropiat de normal<sup>6</sup>.  
as more close to normal.  
“If you already suffer from diabetes, then it is very important that your doctor should help you maintain the blood glucose level as close to the normal level as possible.”
- b. The question complementizer (the Q/wh morpheme), introducing direct or indirect questions.  
Ai pus deja ochii pe casa visurilor tale?<sup>7</sup>  
Have.2SG. put already eyes on house.DEF dreams.DEF.GEN your?  
“Have you already found your dream house?”

- c. psychological-emotive verbs such as *surprise, amaze, etc.*; directive-attitude verbs: *want, insist, suggest, allow*)

Sunt surprins că deja ați hotărât să vă mutați.  
(I)Am surprised that already have.2PL decided to REFL move.  
“I am surprised that you have already decided to move house.”

Sunt uimit că ai ales deja membrii  
(I)Am amazed that have.2SG chosen already members.DEF  
echipei.  
team.DEF.GEN

“I am surprised that you have already chosen the members of the new team.”

- d. Quantifiers: *few, little, every, only* DP, quantifying adverbials: *whenever, wherever*.

Puțini părinți se gândesc deja la carierele  
Few parents REFL think already about careers.DEF  
copiilor.  
children.DEF.GEN

“Few parents are already thinking about their children’s careers.”

- e. incorporated negatives: *doubt, dislike, unlikely, etc.*

Mă îndoiesc că a aflat deja vestea.  
REFL doubt.1SG that have.3SG found out already news.DEF  
“I doubt that s/he has already found out the news.”

- f. negative frequency adverbs: *seldom, rarely*.

Rareori suntem mulțumiți de planurile deja  
Rarely are.1PL content of plans.DEF already  
realizate.  
accomplished.

“Rarely are we content with the things we’ve already accomplished.”

Rareori suntem mulțumiți de ceea ce avem deja.  
Rarely are.1PL content of that-what have.1PL. already.  
“Rarely are we content with what we already have.”

- g. comparatives/superlatives, etc.

Cele mai îngrozitoare scenarii îmi sunt deja  
 DEF most terrifying scenarios CL.DAT.1SG are already  
 familiare.  
 familiar.  
 “I’m already familiar with the most terrifying scenarios.”

- h. Relative clauses headed by indefinite or negative determiners like *any, no, every, a*.

Niciun copil care are deja probleme la școală nu  
 No child who has already problems at school not  
 va îndrăzni să comenteze.  
 will dare SUBJ comment.  
 “No child who is already in trouble at school will dare make any  
 comments.”

- (20) a. The conditional complementizer *if*, more generally  
 hypothetical clauses.

Dacă încă ești supărată pe mine înseamnă că nu  
 If still are angry on me means that not  
 a fost o simplă neînțelegere.  
 has been a simple misunderstanding.  
 “If you are still angry with me it means that it wasn’t just a simple  
 misunderstanding.”

- b. The question complementizer (the Q/wh morpheme), introducing  
 direct or indirect questions.

Încă ești la serviciu?  
 Still are.2SG at work  
 “Are you still at work?”

- c. psychological-emotive verbs such as *surprise, amaze*, etc.; directive-  
 attitude verbs: *want, insist, suggest, allow*)

Sunt surprins că încă mai păstrezi acele poze.  
 Am surprised that still more keep those photos.  
 “I am surprised that you are still holding on to those photos.”

Sunt uimit că încă mai crede în Moș Crăciun.  
 Am amazed that still more believes in Father Christmas.  
 “I am amazed that s/he still believes in Father Christmas.”

- d. Quantifiers: *few, little, every, only* DP, quantifying adverbials: *whenever, wherever*.

Puțini copii încă salută pe oricine le iese  
 Few children still greet DOM everyone CL.ACC.3PL comes  
 în cale.  
 in way.  
 “Few children still greet everyone that comes their way.”

- e. incorporated negatives: *doubt, dislike, unlikely, etc.*

Mă îndoiesc că încă o mai iubește.  
 REFL doubt that still CL.ACC.3SG.FEM more loves.  
 “I doubt that he still loves her.”

- f. negative frequency adverbs: *seldom, rarely*.

Rareori suntem încă plini de speranță că viitorul va  
 Rarely are.2PL still full of hope that future.DEF will  
 fi nemaipomenit când vedem atâta suferință la tot pasul.  
 be great when see.1PL so pain at all step.DEF  
 “Rarely are we still hopeful for a bright future when we see so much  
 pain everywhere we look around.”

Rareori suntem mulțumiți că încă ne mai putem bucura  
 Rarely are.2PL content that still REFL more can enjoy  
 de un simplu răsărit de soare.  
 of a simple rise of sun.  
 “Rarely are we happy that we can still enjoy watching the sun rise.”

- g. comparatives/superlatives, etc.

Cele mai bune prăjituri încă sunt cele făcute de mama mea.  
 DEF more good cakes still are those made by mother.DEF my.  
 “The best cakes are still the ones made by my mum.”

- h. Relative clauses headed by indefinite or negative determiners like *any, no, every, a*.

Niciun părinte care încă mai crede în acest sistem  
 No parent that still more believes in this system  
 de educație nu va proceda așa.  
 of education not will act this way.

“No parent that still believes in this education system will take such decisions.”

- (21) a. The conditional complementizer *if*, more generally hypothetical clauses.

Dacă doar te simți nesigur vom găsi o rezolvare.  
 If only REFL feel insecure will.1PL find a solution.

“If it’s that you only feel insecure, we will find a solution.”

- b. The question complementizer (the Q/wh morpheme), introducing direct or indirect questions.

Ești doar trist sau ești deprimat?  
 Are.2SG only sad or are.2SG depressed?

“Are you only sad or are you depressed?”

- c. psychological-emotive verbs such as *surprise, amaze, etc.*; directive-attitude verbs: *want, insist, suggest, allow*)

Sunt surprins că ai înțeles doar atât din toată  
 Am surprised that have.2SG understood only that from all  
 conversația.  
 conversation.DEF

conversation.DEF

“I am surprised that you only understood that much from the discussion we had.”

Sunt uimit că ți-au rămas doar atâția bani.  
 Am amazed that CL.DAT.2SG-have.2PL left only that money.

“I am amazed that you only have that little money left.”

- d. Quantifiers: *few, little, every, only* DP, quantifying adverbials: *whenever, wherever*.

Puțini bebeluși dorm doar o oră pe noapte.



Few babies sleep only an hour on night.  
 “Few babies sleep only one hour a night.”

- e. incorporated negatives: *doubt, dislike, unlikely*, etc.

Mă îndoiesc că are doar 5 lei.  
 REFL doubt that have.3SG only 5 ron.  
 “I doubt that he only has 5 ron.”

- f. negative frequency adverbs: *seldom, rarely*.

Rareori ai doar un singur motiv de tristețe.  
 Rarely have.2SG only one single reason of sadness.  
 “Rarely has one got only one single reason to be sad.”

- g. comparatives/superlatives, etc.

Cele mai scumpe haine pot fi doar cele mai  
 DEF most expensive clothes can be only DEF most  
 populare fără vreun motiv aparte.  
 popular without any reason particular.  
 “The most expensive clothes can be only those that are the most  
 popular for no particular reason.”

- h. Relative clauses headed by indefinite or negative determiners like  
*any, no, every, a*.

Niciun copil care are doar o temă de făcut  
 No child that has only a homework of done  
 nu va cere să lucreze în plus.  
 not will ask to work in extra.  
 “No child that has only little homework to do will ask to be given  
 extra work.”

Given the hierarchy strength in Figure 1 that we will represent as a table in the following figure, we can say that at least for now, as it turns out from the felicitousness of the previous examples in the previous section, it would look like *deja* (‘already’), *încă* (‘still’) and *doar* (‘only’) are strong positive polarity items. If that were so it means that the previously mentioned polarity sensitive items would not be felicitously licensed in downward

entailing or anti-additive contexts, since the conditions are downwards applicable, as previously discussed.

<b>Negation/ Operators</b>	<b>Superstrong</b>	<b>Strong</b>	<b>Weak</b>	<b>Superweak</b>
Non-veridical	*	√	√	√
Downward Entailing ( <i>At most n</i> )	*	*	√	√
Anti-Additive ( <i>without, refuse</i> )	*	*	*	√
Anti-morphic (negation <i>not</i> )	*	*	*	*

**Table 1: PPIs in negative contexts**

Most likely all of the previous expressions can be felicitously licensed in downward entailing contexts so at most they are weak positive polarity items. But this issue remains to be investigated in a future article, along with the licensing mechanism at a larger scale.

## 4. Conclusions

Many proposals in the literature aimed at providing a classification of polarity items along semantic lines based on the negative environments where they are felicitously licensed. Zwarts (1981) proposed a two-tier classification dividing negative polarity sensitive items into strong and weak, the latter being felicitous in all downward entailing environments. As Hoeksema (2012: 4) correctly points out, there are two great advantages of the classification proposed by Zwarts (1981): one is that it makes use of only one semantic notion, namely downward-entailingness. “Second, the classification is based on cumulative requirements for each higher level of negativity” (Hoeksema 2012: 4).

The present paper only focused on three PPIs and their occurrence in non-veridical contexts. As the examples indicated, all of the three polarity sensitive items are felicitous in non-veridical contexts, finding which means that, according to the diagram proposed by Hoeksema (2012), they are at most strong positive polarity items. The only possibility for them to have been superstrong PPIs would have been if they were ungrammatical in all of the previous examples, but that was not so. The aim of the article was not

to investigate the distribution of these three PSIs in other negative contexts but since the items were grammatical in contexts with *puṭini* (few), which is described as downward-entailing in the literature, the intuition is that these three PSIs are felicitous in downward-entailing contexts as well. In that case, they go further down in the classification at they will be at most weak PSIs.

Given that in the conclusion we mentioned that our three PPIs are likely to also be grammatical in downward-entailing contexts and taking a closer look at what the previous table indicates, it looks like downward-entailing expressions are also non-veridical. This is where an interesting thought may be brought into discussion. Based on Giannakidou's (1998) work, we concluded that every sentence where the speaker is not fully committed to the truth of the sentence, where the truth of the sentence is not fully asserted or presupposed, is non-veridical. Reasoning along the same line, we label a context as nonveridical if whenever  $Fp$  is true  $p$  may or may not be true. So, it could be tricky to say that downward-entailingness can be subsumed under non-veridicality. In an interrogative of the type mentioned under (19b), *Ai pus deja ochii pe casa visurilor tale?* ("Have you already found your dream house?"), it is quite difficult to claim that the superset – *You have found your dream house* – is true or that the subset – *You have found your Victorian dream house* – is true. As it is quite difficult to clearly claim which of the previous options is true, maybe non-veridicality would best be understood as non-monotonic.

Another point worth mentioning and worth exploring in a future article is the point made by Allen (2006) who claims that negated sentences may alternatively be evaluated as factual statements in that we can treat NOT  $p$ , *it is not the case that p* where we treat negative sentences as anti-veridical, as *it is the case that NOT p*, where a sentence is treated as asserting the truth of NOT  $p$ . It remains to be seen what implications such claims have in the licensing of polarity sensitive items, especially for PPIs.

## References

- Allen, Keith. 2006. "Clause-type, primary illocution, and mood-like operators in English". *Language Sciences*, 28 (1): 1-50.
- Giannakidou, Anastasia. 1998. *Polarity Sensitivity as (Non)veridical Dependency*. Amsterdam-Philadelphia: John Benjamins.
- Giannakidou, Anastasia, and Lisa Lai-Shen Cheng. 2006. "(In)Definiteness, Polarity, and the Role of wh-morphology in Free Choice". *Journal of Semantics* 23 (2):135-183.

- Giannakidou, Anastasia. 2011. "Positive polarity items and negative polarity items: variation, licensing, and compositionality". In *Semantics: An International Handbook of Natural Language Meaning*, edited by Claudia Maienborn, Klaus von Stechow, and Paul Portner, 1660-1712. Berlin: Mouton de Gruyter.
- Hoeksema, Jack. 2012. "On the Natural History of Negative Polarity Items". *Linguistic Analysis* 38 (1-2): 3-33.
- Horn, Laurence. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Klima, Edward. 1964. "Negation in English". In *The structure of Language*, edited by Jerry A. Fodor, and Jerrold J. Katz, 246-323. Englewood Cliffs, NJ: Prentice-Hall.
- Ladusaw, William. 1979. *Polarity Sensitivity as Inherent Scope Relations*. PhD diss., University of Texas at Austin.
- Linebarger, Marcia. 1987. "Negative polarity and Grammatical Representation". *Linguistics and Philosophy* 10 (3): 325-87.
- Montague, Richard. 1969. "On the nature of certain philosophical entities". *The Monist* 53: 159-194. 1974. Reprinted in *Formal Philosophy: Selected papers of Richard Montague*, edited by Richmond H. Thomason, 148-187. New Haven: Yale University Press.
- Szabolcsi, Anna. 2010. *Quantification*. Cambridge: Cambridge University Press.
- Van der Wouden, Ton. 1997. *Negative Contexts: Collocation, Polarity and Multiple Negation*. London: Routledge.
- Zwarts, Frans. 1981. "Negatief Polaire Uitdrukkingen I". *GLOT* 4: 35-133.
- Zwarts, Frans: 1995. "Nonveridical contexts". *Linguistic Analysis* 25: 286-312.

## Notes

<sup>1</sup> Such sentences are not of interest for the present paper. For further details on emphatic denial, see Horn (1989). PPIs can occur in the immediate scope of negation if this is understood as an emphatic denial of a previously stated sentence. This is exemplified with the previously mentioned example:

Sylvia has *something* to write for the English lesson.

Wrong! Sylvia DOES NOT have *something* to write for the English lesson.

<sup>2</sup> See van der Wouden (1997) where antimorphic is treated as  $f(X \cap Y) = f(X) \cup f(Y)$ .

<sup>3</sup> See van der Wouden (1997): Anti-additive  $f(X \cup Y) = f(X) \cap f(Y)$  – nobody, never, nothing.

<sup>4</sup> The formal definition of veridicality views the context as a propositional operator. By veridicality we understand the *truth of a proposition in a situation or in a context*.

---

a. A propositional operator  $F$  is veridical iff  $Fp$  entails  $p$ :  $Fp \rightarrow p$ ; otherwise  $F$  is nonveridical.

b. Additionally, a nonveridical operator  $F$  is antiveridical iff  $Fp$  entails not  $p$ :  $Fp \rightarrow \neg p$ .

(Giannakidou, 2011) In other words, a non-veridical operator  $F$  is anti-veridical iff whenever  $Fp$  is true,  $p$  is not true.

<sup>5</sup> “In the definition above, I use of a multimodel system where truth of a proposition is evaluated with respect to an individual’s epistemic model,  $ME(x)$ .  $ME(x)$  is a set of worlds representing the epistemic status of the individual  $x$ ” (Giannakidou, 2011:12)

“For  $\alpha$  believes that  $p$  to be true, it must be the case that  $\alpha$ , the main clause subject, is committed to the truth of the embedded proposition  $p$ . Though the speaker might disagree, a prerequisite for  $p$  to be true in (48) is that Jacob’s epistemic model (i.e. the set of worlds compatible with what Jacob believes) be a subset of the worlds where  $p$  is true:  $ME(\text{Jacob}) \subseteq p$ . The speaker may believe or even know that what Jacob believes is false, but this is irrelevant for Jacob’s beliefs.

(2) [[ Jacob believes that Ariadne loves Paul]]  $c = 1$  iff

$\forall w [w \in ME(\text{Jacob}) \rightarrow w \in \lambda w'. \text{Ariadne loves Paulin } w']$ ” (Giannakidou, 2011:14)

<sup>6</sup> <https://www.cardioportal.ro/pacienti/factori-de-risc-cardiovascular/diabetul-zaharat/ce-pot-sa-fac-daca-am-deja-diabet-zaharat/>

<sup>7</sup> OTP BANK, <https://www.facebook.com/watch/?v=2307875202826336>

# CHAPTER EIGHT

## TWO TYPES OF QUANTITY RELATIVE SUPERLATIVES

ION GIURGEA

I discuss superlative *most* in a number of languages, showing that some languages show sensitivity to c-command by the correlate: in some of these languages, in non-c-commanded positions the superlative *most* is differently marked (usually by the presence of the definite article), in others it is completely disallowed, other constructions being used instead (*the largest part, the majority, the largest number*). I explain this contrast by the existence of a c-command requirement for -EST raising out of the DP. For the forms used when the requirement is not satisfied, I propose an in-situ derivation of the relative reading, based on association with focus, as suggested in Heim (1999). An important consequence for the theory of superlatives is that relative readings are obtained both via -EST raising and via the in-situ strategy (whereas in Heim 1999 these derivations are presented as competing analyses).

Keywords: superlatives, relative readings, quantitatives

### 1. Introduction

This paper reports and tries to explain an interesting split in the use of quantity relative superlatives, found in a number of languages.

Starting with Szabolcsi (1986) and Heim (1985, 1999), a distinction is made between two readings of superlatives, the *absolute* and *relative* (or *comparative*) reading<sup>1</sup>. It is generally agreed that the semantics of superlatives involves choosing the maximal degree from a set of degrees. Absolute and relative readings differ in the way this set is built. In the absolute reading, the compared degrees are associated with all the entities in the denotation of the NP to which the superlative is attached (with the

usual domain restrictions) – in other words, the *comparison class* is provided by DP-internal material only, see (1) a. In the relative reading, the compared degrees are described by a property provided by the clause in which the superlative occurs. The compared degrees do not differ just with respect to the entity to which the property applies (the various mountains in example (1)), but also with respect to another constituent in the clause, the so-called *correlate* (in (1)b, the correlate is the subject; each of the compared degrees is associated with a different choice of the subject, e.g., John Mary, Alice, Christian, etc.):

- (1) John climbed the highest mountain  
 (a) absolute: John climbed Mount Everest (or the highest mountain in a contextually restricted set of mountains)  
 Compared degrees: {d:  $\exists x$ . x is a mountain and x is d-high}  
 (b) relative: John climbed a mountain higher than the mountains climbed by everybody else  
 Compared degrees: {d:  $\exists x$ .  $x \in \{\text{John}, a, b, \dots\}$  and x climbed a d-high mountain}

Szabolcsi (1986) observed that quantity superlatives only have relative readings. As in the absolute reading the comparison class C is in principle established based on the NP property alone, we get indeed absurd readings for absolute superlatives: since the sum of all pluralities/portions of stuff is in the denotation of the NP, the element with the maximal degree is this maximal sum itself, which means that *most* in the absolute superlative reading would mean *all*. For *fewest*, as there is no single entity in the denotation of a plural or mass term that has a minimal cardinality or measure, the superlative description ‘fewest N’ will never be satisfied.

Hackl (2009), introducing an amendment into the semantics of superlatives, identified the missing absolute reading of *most* with the proportional reading<sup>ii</sup>. However, further research has shown that this analysis is untenable on empirical grounds, because many languages have relative *most* but no proportional *most* (see Dobrovie-Sorin 2013, 2015, Dobrovie-Sorin & Giurgea (2021), Coppock et al. 2017, Coppock 2019).

During the research on the way of expressing majority quantification and the possible interpretations of MOST across languages that Carmen Dobrovie-Sorin and I carried out (see Dobrovie-Sorin & Giurgea 2021.), we discovered an unexpected contrast in the way of expressing the *relative* superlative of MANY/MUCH. In some languages, superlatives not c-commanded by the correlate behave differently than those c-commanded

by the correlate, whereas in other languages no such contrast is visible. This paper tries to offer an explanation for this unexpected contrast.

In section 2, I will present the facts. In section 3, I will briefly revise the main types of analyses of relative superlatives. In section 4, I will argue that the contrast in the way of expressing relative superlative readings of quantity depending on the position of the correlate, which appears to be limited to indefinite superlatives, follows from constraints on -EST raising. The consequence of this explanation, developed in section 5, is that the relative superlatives found in non-c-commanded positions, in the languages that have the contrast, do not rely on -EST raising. The general conclusion is that both mechanisms proposed in the literature for deriving relative readings, namely -EST raising and contextual restriction of the comparison class, are actually used by language.

## 2. The data: crosslinguistic variation with superlative MOST

The literature on relative superlatives (see Szabolcsi 1986, Heim 1999, Farkas & Kiss 2000) has established that the correlate is usually either the focus or a variable bound by a *wh*-operator (interrogative or relative):

- (2) a. John gave ALICE the most expensive present  
: for all *x* different from Alice, John gave *x* a less expensive present than he gave Alice
- b. JOHN gave Alice the most expensive present  
: for all *x* different from John, *x* gave Alice a less expensive present than John gave her
- c. Who gave Alice the most expensive present?  
: who is *x* such that for all *y* different from *x*, *y* gave Alice a less expensive present than *x* gave Alice
- d. John is the person who gave Alice the most expensive present  
: for all *x* different from John, *x* gave Alice a less expensive present than John gave her

In the questionnaires used in Dobrovie-Sorin & Giurgea (2021)<sup>iii</sup>, we checked for relative uses of the superlative of MANY/MUCH by using both foci and *wh*-words as correlates. In what follows, focus will be marked by using capitals and the correlate will be underlined:



- (3) a. JOHN has the most friends (of all my colleagues)  
 b. JOHN has the most money (of all my colleagues)  
 c. Who has the most friends?  
 d. JOHN speaks the most.
- (4) The most/largest number of immigrants come from INDIA (compared to other countries)

Notice that in (3) the correlate c-commands the superlative, whereas in (4), the superlative c-commands the correlate. What we observed is that, whereas certain languages use the same form for the superlative in (3) and (4), in a number of other languages the forms used for rendering (3) and (4) differ: in some of them, (4) differs from (3) by the use of an article; in others, the superlative form of MANY/MUCH (which I notate as MOST, with capitals) is marginal or excluded in examples of the type in (4), the relevant meaning being expressed via a different construction. We discovered that English itself is of this latter type, therefore we adapted the example by adding the variant ‘the largest number of...’. I label the types of languages we discovered ‘type A’ and ‘type B’ (with the further division into ‘type B1’ and ‘B2’), defined as follows:

- (5) a. Type A languages: no difference between c-commanded and non-c-commanded position  
 b. Type B languages: difference between c-commanded and non-c-commanded position  
 B1: the non-c-commanded position differs from the c-commanded position by the use of an article  
 B2: MOST is disallowed in non-c-commanded positions; the relevant meaning is conveyed by expressions of the type THE LARGEST PART, THE LARGEST NUMBER, or even THE MAJORITY

**Type A languages** include Romanian, Greek, Slavic languages, German, Icelandic, Persian, Hungarian, Hebrew, Japanese. Let me cite here Hebrew as an example, which is particularly conclusive because Hebrew does not use MOST for the proportional reading, so the use of MOST in (6)b (the counterpart of (4)) cannot be due to a confusion with the proportional reading.<sup>iv</sup>

- (6) a. le-Dan yef **haxi harbe** xaverim.  
 DAT-Dan EXIST SUP many friends  
 ‘Dan has the most friends.’
- b. **haxi harbe** mehagrim ba’-im me-hodu.  
 SUP many immigrants come.PTPL-MPL from-India  
 ‘The largest number of immigrants come from India’

Other languages of type A where MOST does not have a proportional reading are Japanese and Slavic languages. Among the latter, we checked Bulgarian. For other Slavic languages, the absence of a proportional reading of MOST as well as the use of MOST in non-c-commanded contexts (examples of the type in (4)) have been reported in Živanović (2007), who has data from Slovenian, Czech, Polish, Serbian, and Macedonian.

In **languages of the type B1**, MOST in the non-c-commanded environment in (4) differs from the MOST used in c-commanded environments by the presence of an article. In Armenian, this is the definite article, see (7); in Basque, which does not have a definite article, it is a general ‘argumental’ article (typically used with nouns in argumental positions), see (8):

- (7) a. Ov uni **amena-šat** ěnkerner?  
 who has SUP-many friends  
 ‘Who has the most friends?’
- b. **Amena-šat** emigrantner-ě galis en Hndkastan-ic.  
 SUP-many immigrants-the coming are India-ABL  
 ‘The most (largest number of) immigrants come from India’
- (8) a. Nork du [lagun **gehien**] nere kolegen artean?  
 who.ERG has friend most my colleague.GEN among  
 ‘Who has the most friends (of all my colleagues)?’
- b. [Inmigrante **gehien-a-k**] Indiatik heldu dira  
 immigrant most-ART-PL India.from come are  
 ‘The largest number of immigrants come from India.’

Interestingly, the same MOST+ART form is used in Basque for the proportional reading. In Armenian, MOST lacks the proportional reading, which is conveyed by expressions of the type THE LARGEST PART. Note furthermore that in Basque the article is optional in the case of

cardinals, and its use correlates with definiteness, according to Artiagoitia (2002):

- (9) a. [Hiru tren] heldu dira. (Artiagoitia 2002:84)  
 three train arrive AUX  
 ‘Three trains arrived.’  
 b. [hiru tren-a-k] heldu dira.  
 three train-ART-PL arrive AUX  
 ‘The three trains arrived.’

Swedish resembles Basque in that the use of the (definite) article with MOST is associated to the proportional reading (see Coppock & Josefson 2015, Coppock 2019):<sup>v</sup>

- (10) a. Gloria har besökt **de flesta** kontinenterna.  
 Gloria has visited the.PL most continents-the  
 ‘Gloria has visited most continents (i.e. more than half)’  
 b. Gloria har besökt **flest** kontinenter.  
 Gloria has visited most continents  
 ‘Gloria visited the most continents (i.e., more than the others)’  
 (Coppock & Josefson 2015: ex. 3-4)

This distinction is weakened in the non-c-commanded position. Although the articleless variant is still acceptable, one of our informants used the definite article, showing the following contrast:

- (11) a. Vem har **flest** vänner?  
 who has most friends  
 b. **De flesta** invandrarna kommer från Indien.  
 the most immigrants.DEF come from India

Other speakers report, however, that the use the article triggers a proportional interpretation in (12)b:

- (12) a. **Flest** barn {är födda/föds} i juli  
 most children are born in July  
 ‘The largest number of children are born in July’  
 b. **De flesta** barnen {är födda/föds} i juli  
 the most children.DEF are born in July  
 ‘Most / %The largest number of the children are born in July’

This suggests that Swedish oscillates between type A and type B1.<sup>vi</sup>

For Norwegian, our only informant provided a translation with THE MOST (*de fleste*) for the immigrants example in (4). Moreover, even for the babies example in (12), where a proportional reading is unlikely, he used the definite article, see (13), in contrast with the non-c-commanded position in (14):

- (13) **De fleste** babier er født i juli.  
 the most babies are born in July  
 ‘The most/largest number of babies are born in July.’
- (14) Hvem har **flest** venner?  
 who has most friends  
 ‘Who has the most friends?’

Let us now turn to **type B2 languages**. Here, the most striking case is English: although *the most*, with what appears to be a definite article, is used in the c-commanded position, *the most* cannot be used in the non-c-commanded position, see (16)a. In order to render the desired interpretation, informants suggested the variant in (16)b, with *the largest number*. Even *most*, which normally triggers a proportional interpretation, has been accepted as a possible relative superlative by some informants (three out of seven):

- (15) Who has **the most** friends?
- (16) a. ?? **The most** immigrants come from India.  
 b. **The largest number** of immigrants come from India.  
 c. % **Most of** the immigrants come from India.

Other languages of type B2 are French, Italian, Catalan, Spanish, Portuguese, and Albanian. These languages lack dedicated superlative morphology, using comparatives embedded in definite DPs to convey superlative readings. A special case is French, where the definite article *le* came to function as a superlative marker with postnominal, adverbial and predicative superlatives (see Loccioni 2018 for evidence that this ‘article’ is not a ‘nominal’ determiner, but is part of the Degree projection of the adjective/adverb). I show below two cases where the fact that French *le* is not a D is indisputable: adverbial superlatives, see (16), and predicative superlatives which do not rely on N ellipsis (do not have a postcopular DP, but rather a DegP/AP predicate), see (17); in this example, the fact

that *le* is not an article (a D) is shown by the lack of agreement (*le* is a default form, if it were an article we would have expected the feminine form *la*) and by the possibility of using something else than the subject as variable element across the descriptions of the compared degrees (in (17), the comparison is between circumstances when a mother is happy, not between the mother and other persons; such an interpretation is impossible if the superlative is embedded in a postcopular DP, see Loccioni 2018:71-73):<sup>vii</sup>

- (16) Elle parle **le plus fort** (de tous).  
 she speaks the more loud of all  
 ‘She speaks/is speaking the loudest (of all).’
- (17) C’est au milieu de ses enfants qu’une mère est **le plus heureuse**.  
 it is in-the middle of her children that a mother(F) is the.MSG  
 more happy.FSG  
 ‘It is among her children that a mother is happiest.’  
 (Grevisse 2008: 1229-30)

In French, *le plus* lit. ‘the more’ occurs as a quantity superlative in c-commanded positions (see (18)), but is judged highly marginal in non c-commanded positions, see (19):

- (18) a. Qui a **le plus** d’amis?  
 who has the more of friends  
 ‘Who has the most friends?’
- b. Guillaume a vu **le plus** de pays.  
 Guillaume has seen the more of countries  
 ‘Guillaume has seen the most countries.’  
 (Loccioni 2018:185)
- (19) ?? **Le plus** de migrants viennent de l’Inde.  
 the more of immigrants come from India

In Italian and Ibero-Romance, the article is not used as a superlative marker with quantity superlatives (in fact, it only has a superlative-like import in a special modal construction of the type *the best/most (...)* possible, see Loccioni 2018). Neither is embedding of MORE in a definite DP used to convey the meaning of relative MOST. As a consequence, MORE alone functions sometimes as a superlative, but under very

restricted conditions: the correlate must be bound by a *wh*-word; therefore mere prosodic focus by itself does not license the superlative, a cleft construction is needed, see (20)a,b,e:

- (20) a. En Joan és el que té **més** diners/amics. (Cat.)  
 ART Joan is the that has more money/friends  
 ‘Joan has the most money/friends.’
- b. Juan es el que {tiene **más** amigos / habla más}. (Sp.)  
 Juan is the that has more friends speaks most  
 ‘Juan has the most friends/speaks the most.’
- c. Quién tiene **más** amigos? (Sp.)  
 who has more friends  
 ‘Who has the most friends?’
- d. Quem tem **mais** amigos? (Port.)  
 who has more friends  
 ‘Who has the most friends?’
- e. Gianni è quello che ha **più** soldi / amici. (It.)  
 Gianni is the-one that has more money/friends  
 ‘Gianni has the most money/friends.’

Albanian resembles Italian and Ibero-Romance by lacking the use of the article with *më shumë* ‘more many/much’, but does not require a *wh*-word as a correlate, see (21)b, where the use of the PP ‘among all’ disambiguates in favor of the superlative interpretation:

- (21) a. Kush ka **më shumë** shokë?  
 who has more many friends  
 ‘Who has the most friends?’
- b. Xhoni flet **më shumë** nga të gjithë.  
 John speaks more much from all  
 ‘Of all, John speaks the most.’

In all these languages, in order to convey the relative superlative reading in the immigrants-example in (4), the informants used expressions that normally convey the proportional (majority) interpretation: THE LARGEST PART, THE MAJORITY (although they were asked to think of a context where the number of immigrants coming from India is less than 50%); one Spanish informant used a LARGEST NUMBER construction (see (22)d):<sup>viii</sup>

- (22) a. **La plupart** des migrants vient de l'Inde. (Fr.)  
the more-part of-the immigrants comes from the-India
- b. **La maggior parte** dei migranti viene dall'India. (It.)  
the major part of-the immigrants comes from-the India
- c. **La major part** dels immigrants provenen de l'India. (Cat.)  
the major part of-the immigrants come from the India
- d. **La mayor cantidad** de inmigrantes vienen de la India. (Sp.)  
the major quantity of immigrants come from the India
- d'. **La mayoría** de los inmigrantes provienen de la India. (Sp.)  
the majority of the immigrants come from the India
- e. **A maioria** dos imigrantes vem da Índia. (Port.)  
the majority of-the immigrants come(s) from-the India
- f. **Shumica** e emigrantëve vijnë nga India (Alb.)  
majority-the AGR migrants-the.GEN come from India-the  
'The most /largest number of immigrants come from India  
(compared to other countries)'

### 3. Main types of accounts of relative superlatives

The first studies that addressed the absolute vs. relative distinction (Szabolcsi 1986, Heim 1985, 1999) explained it as a difference in scope: whereas in absolute readings, the superlative operator -EST remains inside the DP, taking scope over the NP, in the relative reading, it raises out of the DP and takes clausal scope:

- (1) a. absolute: THE [-EST  $\lambda d.\lambda x.[\text{mountain}(x) \wedge \text{d-high}(x)]$ ]  
b. relative: John [-EST  $\lambda d.\lambda x.[\exists y(\text{mountain}(y) \wedge \text{d-high}(y) \wedge \text{climb}(\text{John},y))]$ ]

Thus, whereas in the absolute reading the comparison class consists of mountains, in the relative reading, the comparison class is a set that contains the correlate (in (1)b, John and other individuals). The denotation of -EST is the same in both readings (it takes a comparison class C, a function R from degrees to properties of entities (see the  $\lambda d.\lambda x$ -expressions in (1)' above), and yields a property of individuals – see  $\lambda x$  in (22)):

- (22)  $[-EST] = \lambda C \lambda R_{\langle d, \langle e, t \rangle \rangle} \lambda x \exists d (R(d)(x) \wedge \forall y ((y \in C \wedge y \neq x) \rightarrow \neg R(d)(y)))$   
 $[-EST] (C)(R)(x)$  is defined iff  
 (i)  $x \in C$   
 (ii)  $\forall y (y \in C \rightarrow \exists d R(d)(x))$   
 (iii)  $\exists y (y \in C \wedge y \neq x)$  (adapted after Heim 1999)

Note that in this analysis, the superlative DP (i.e., the DP containing the superlative) is interpreted as indefinite in the relative reading, in spite of the presence of the definite article. Szabolcsi (1986) showed that it patterns, indeed, with indefinites in certain contexts where definites are not allowed (see (23)-(24)), concluding that it is indeed interpreted as an indefinite, the article being spurious here:

- (23) a. John has the smartest sister  
 b. \* John has the (smart) sister  
 c. John has a (smart) sister
- (24) a. Who did you take the best picture of?  
 b. \*Who did you take the picture of?  
 c. Who did you take a picture of?

Further evidence for raising of -EST was provided by Heim (1999): in intensional contexts, such as (25), there is a relative reading where the degree operator -EST scopes above the modal operator, but the indefinite DP scopes below – this is the so-called ‘upstairs de dicto reading’, see (25)c:

- (25) JOHN wants to climb the highest mountain  
 a. de re reading: John wants to climb a certain mountain, which is higher than the mountains the others want to climb  
 John  $[\lambda x. -EST \lambda d [\exists y. y \text{ is a } d\text{-high mountain} \wedge \text{WANT}(x, x \text{ climbs } y)]]$
- b. ‘downstairs’ de dicto reading: John wants to climb a(ny) mountain that is higher than any of the mountains the others will climb (John wants that the mountain he will climb should turn out to be higher than the mountains climbed by all the others)  
 John  $[\lambda x. \text{WANT}(x, -EST \lambda d [\exists y. y \text{ is a } d\text{-high mountain} \wedge x \text{ climbs } y)]]$



c. ‘upstairs’ de dicto reading: John wants to climb any mountain that has a certain height; this height is greater than the height the others want the mountains they climb to have (de re comparison; de dicto superlative DP)

John [ $\lambda x$ . -EST  $\lambda d$  [WANT ( $x$ ,  $\exists y$ .  $y$  is a  $d$ -high mountain  $\wedge x$  climbs  $y$ )]]

The readings (25)a-b can also be derived if we adopt a non-raising approach to relative superlatives, in which focal alternatives are used to constrain the comparison class: thus, in (25)a, *the highest mountain* can be treated as a specific definite, referring to the highest mountain among the mountains that someone wants to climb. In (25)b, *the highest mountain* can be treated as a non-specific definite, with uniqueness relativized to the possible worlds introduced by the modal: thus, in any world of John’s wishes, *the highest mountain* refers to the unique mountain that is higher than the mountains the others climb in that world. However, the reading in (25)c cannot be derived in this way: whereas the object is non-specific, interpreted de dicto, the comparison itself is de re, out of the scope of the modal. Heim concludes that this reading provides clear evidence for raising of -EST in the clause, independent of the scope of the superlative DP.

However, the consistent use of the definite article with relative superlatives, in various different languages, remains unexplained in the raising analysis. Krasikova (2012), developing a suggestion of Szabolcsi’s, proposes that THE in relative superlatives is part of the superlative DegP, introducing maximalization over degree properties. Krasikova’s account is based on an alternative semantic analysis of DP-external -EST, which was proposed in Heim (1999) as a way of better accounting for the association of relative superlatives with focus. In this analysis, -EST scopes above the correlate and takes a property of degrees (see P in (26)), instead of a relation between entities and degrees, as its second argument, whereas the first argument, the comparison class (C in (26)), is a set of degree properties:

$$(26) \quad \begin{aligned} [-\text{EST}] &= \lambda C_{\langle d, t \rangle} \lambda P_{\langle d, t \rangle} \exists d(P(d) \wedge \forall Q ((Q \in C \wedge Q \neq P) \rightarrow \\ &\quad \neg Q(d))) \text{ defined iff } P \in C \wedge \exists Q(Q \neq P \wedge Q \in C) \\ &\quad \text{(modeled after Heim 1999:(65))} \end{aligned}$$

$$(27) \quad C\text{-EST } \lambda d \text{ [[John}_F \text{ climbed a } d\text{-high mountain]} \sim C]$$

The comparison class is directly provided by focus: the  $\sim$  operator attached to C forces C to be a subset of the focus value of the clause, which is the set of degree properties of the form  $\lambda d$  [x climbed a d-high mountain].

Note that this semantics cannot apply to DP-internal -EST (unless absolute superlatives are taken to involve a hidden relative clause, which neither of these studies claims). Krasikova, indeed, proposes a different semantics for absolute superlatives, in which the superlative operator combines, in turn, with the adjective, with the NP, and yields a property of individuals (the superlative AP is assumed to scope above the NP and below D; no comparison class is used):

$$(28) \quad \llbracket \text{SUP} \rrbracket = \lambda A_{\langle d, \langle e, t \rangle} \lambda P_{\langle e, t \rangle} \lambda x. \exists d [A(d)(x) \wedge \forall y [(P(y) \wedge y \neq x) \rightarrow \neg A(d)(y)]]$$

For relative superlatives, Krasikova modifies Heim's (1999) analysis given in (26) above in order to implement the idea that THE in relative superlatives introduces an iota-operator over degree properties. The superlative semantics is encoded in an operator *max* which applies before THE. However, Krasikova's (2012) analysis of THE as forming a part of the superlative is empirically wrong, as the article can be separated from the superlative in various languages, including English. This is very clear in Romance languages, where relative readings are available for post-nominal superlatives, and yet the definite article occurs in D (see (29), which allows Heim's upstairs de dicto reading). In English, THE can be separated from a relative superlative by a cardinal, which clearly cannot be part of the DegP (see (30), which allows the upstairs de dicto reading).

(29) Gianni vuole scalare [**la** montagna **più alta**]. (It.)  
Gianni wants to-climb the mountain more high  
'Gianni wants to climb the highest mountain.'

(30) John needs to read [**the** [two [[**longest**] books]]]

Due to the problem of the definite article, several studies adopted analyses of relative superlatives where -EST is interpreted inside the DP, which allows the definite article to keep its uniqueness interpretation (see Farkas & Kiss 2000, Sharvit & Stateva 2002, Teodorescu 2009, Coppock & Beaver 2014). Coppock & Beaver (2014) argue that there are other definite DPs which pattern with indefinites with respect to the tests in (23)-(24):

- (31) a. Mary has the only lazy sister  
 b. They have the same father

They argue that THE should be decomposed into a uniqueness filter on properties and a determiner, which can be *iota*, in definites with determined reference, or a mere existential, in definites without determined reference. The latter case comprises the examples in (31) as well as the DPs with relative superlatives in (23)-(24).

Some in-situ analyses of relative readings introduce into the semantic representation a relation that pairs the compared entities (which belong to the NP denotation) with the elements of the set to which the correlate belongs (the so-called ‘contrast set’), see Farkas & Kiss (2000), Coppock & Beaver (2014). Other analyses consider that relative readings are obtained by putting further contextual restrictions on the comparison class (Sharvit & Stateva 2002, Teodorescu 2009, Pancheva & Tomaszewicz 2012). Heim (1999) considered an in-situ analysis of this type, where the superlative DP is QR-ed to a position above the focus operator, and the comparison class is restricted to be a subset of the union of the focus alternatives of the sister of the DP, as explained in (32):

- (32) a. [the [C-est [d-high mountain]]]  $\lambda x$ . [ [John<sub>F</sub> climbed x]  $\sim$ S]  
 b.  $S \subseteq \{P: \exists y[P=\lambda x. y \text{ climbed } x]\}$  (the set of focal alternatives)  
 c.  $C = \cup S$  (association with focus)  
 d.  $b+c \Rightarrow C \subseteq \{x: \exists y. y \text{ climbed } x\}$

Pancheva & Tomaszewicz (2012) argue that both the raising and in-situ strategies are at work in relative readings, but, crucially, superlatives embedded in definite DPs are always in-situ. Evidence for this proposal comes from the existence of a ban on DP-internal correlates that occurs in definite superlatives, but is suspended in superlatives embedded in bare nominals, which are possible in Polish and Bulgarian (adopting the in-situ strategy based on association with focus, exemplified in (32), they demonstrate that DP-internal correlates cannot be derived by this strategy):

- (33) a. \* John has [the best albums by U2]  
 b. \* Ivan ima [naj-dobri-te albumi na U2] (Bulg.)  
 Ivan has SUP-good-the albums by U2

- c. Ivan ima [naj-dobri albumi na U2] (Bulg.)  
 Ivan has SUP-good albums by U2  
 ‘Among the albums Jan has, those of U2 are better than  
 those of other artists’  
 (Pancheva & Tomaszewicz 2012: ex. 19,15a)

Upstairs de dicto readings remain hard to derive under in-situ approaches.<sup>ix</sup>

Bumford (2017) proposed a raising analysis which accounts for the use of definiteness marking on the superlative DP. Like Coppock & Beaver (2014), he decomposes the article into an existential and a uniqueness component, but he assumes that uniqueness is evaluated on variable assignments, above the DP, at the level where -EST is evaluated. The superlative itself is analyzed as a filter on variable assignments. As this semantics does not make use of degrees, Heim’s upstairs de dicto reading cannot be accounted for.

In sum, the debate between in-situ and raising analyses of relative superlatives in definite DPs remains unsettled.

#### 4. Accounting for the behavior of quantity superlatives

In section 2 we have seen that in most languages belonging to type B, the superlative in c-commanded position occurs in a DP that lacks an article (see Basque, Armenian, Italian, Ibero-Romance, Albanian, Swedish, and Norwegian), whereas in non-c-commanded position, the superlative occurs in a definite DP (see Basque, Armenian, and Norwegian, as well as the LARGEST-PART and MAJORITY-expressions found in Romance languages and Albanian, which are all definite; in Swedish, this is subject to speaker variation). The only exceptions are French and English, but there is evidence that even these exceptions are only apparent. For French, this was already anticipated in our discussion: as *le plus* occurs in adverbial and predicative superlatives (see ex. (16)-(17)), it is reasonable to assume that in the DPs illustrated in (18)-(19), *le plus* forms a constituent in which *le* is a superlative marker, which makes the whole DP indefinite. As for English, Wilson (2018) has argued on independent grounds (based on the availability of NP-internal correlates, see (34)) that *the* in *the most N* is not the definite article of the entire DP, but rather forms a constituent with the superlative (the structure she proposes is given in (35), where the head  $\text{Mon}^0$ , taken from Schwarzschild 2006, is the functional projection that introduces quantity modification):

- (34) He ate [the most CHOCOLATE mini-cupcakes]  
 = He ate more chocolate mini-cupcakes than he ate of any other  
 type (Wilson 2018: 26, ex. 31)
- (35) [<sub>DP</sub>  $\emptyset_{\exists}$  [<sub>MonP</sub> [<sub>DP</sub> the most AMOUNT] [<sub>Mon</sub><sup>0</sup> [<sub>NP</sub> chocolate mini-  
 cupcakes]]]]]

Now, from the discussion in section 3 it should be clear that the raising analysis is the only option if the superlative DP is indefinite (an in-situ superlative triggers uniqueness, therefore the definite article is required in languages that mark uniqueness by the definite article). As all quantity superlatives of type B have been argued to be indefinite, we may conclude that the distributional constraint on quantity superlatives reflects a constraint on raising -EST:

- (36) In languages of type B, raising -EST requires c-commanding by the correlate

In fact, a similar constraint was proposed by Farkas & Kiss (2000) for relative superlatives in general, based on the data in (37) (which, crucially, involve *quantity* superlatives); the requirement is formulated in terms of m-command rather than c-command in order to account for (37)c, where the correlate and the superlative DP mutually m-command each other:

- (37) a. John received the fewest votes.  
 b. \* The fewest voters voted for John.  
 (Farkas & Kiss 2000:427, ex. 24)  
 c. Voters cast the fewest votes for John.

The French examples in (38) show that it is not the subject position per se that rules out superlative MOST, but rather the absence of a c-commanding correlate:

- (38) a. C'est sur Napoléon que le plus de livres a/ont été  
 it is on Napoleon that the more of books has/have been  
 écrit(s).  
 written(PL)  
 'It's on Napoleon that the most books have been written.'  
 b. Où est-ce que le plus d'or a été trouvé ?  
 where is it that the more of gold has been found  
 'Where was the most gold found?'

- c. C'est de l'Inde que viennent **le plus d'immigrants**.  
 it is from the India that come the more of immigrants  
 'It's from India that the most immigrants come.'

This indicates that, at least in languages of type B, -EST raising out of the DP is not a freely available LF-operator movement, but obeys syntactic constraints, requiring an attractor of a special type (wh-operator, focus) that c-commands -EST in overt syntax. We may assume that an Agree relation is established between this higher operator and -EST, which triggers post-spell-out movement.

Now, it appears that this constraint is not universal. In Swedish, speakers accepted an indefinite superlative DP in a non-c-commanded position, as in (39)a below, which also shows Heim's upstairs *de dicto* reading. Note that a definite superlative is not felicitous with this reading, as shown in (39)b (the example can only be interpreted as referring to a specific set of pages that have to be written, of which the largest part is for the literature course):

- (39) a. **Flest sidor** ska skrivas i litteraturkursen.  
 most pages must write.PAS in literature-course.DEF  
 'The largest number of pages must be written for the literature course'  
 b. **#De flesta sidorna** ska skrivas i litteraturkursen.  
 the most pages.DEF must write.PAS in literature-course.DEF

The same holds for Bulgarian and Icelandic, where the immigrants-examples were rendered with an indefinite DP by our informants (recall that an indefinite superlative necessarily involves -EST raising):

- (40) a. **Naj-mnogo imigranti** idvat ot India. (Bg.)  
 SUPERL-many immigrants come from India  
 'The largest number of immigrants comes from India'  
 b. **Flestir innflytjendur** eru frá Indlandi. (Ice.)  
 most immigrants are from India  
 'The largest number of immigrants are from India.'

This indicates that the absence of the split in the expression of quantity superlatives in type A languages is due, at least in part, to the absence of an *overt* c-command requirement on -EST raising (we may assume that the c-command requirement is fed by LF-raising of the focus).

Some type A languages have the definite article across the board (see German, Hungarian, Greek). In German and Hungarian, the article can also occur with adverbial superlatives (e.g., Germ. *am besten* ‘at-the best’, Hung. *a legjobban* ‘the best’), so it is possible that this article is part of the superlative DP in the case of quantity superlatives, as has been argued for English by Wilson (2018). This does not seem to apply to Greek, however, so at least for this language we cannot be sure that quantity superlatives are formally indefinite.

In Romanian, which is a type-A language, it is indisputable that the strong definite article *cel* has evolved into a superlative marker when combined with the comparative degree head *mai* (see Cornilescu 2007, Giurgea 2013 a, b, Cornilescu & Giurgea 2013). When occurring DP-initially, superlatives may mark the entire DP as definite (presumably by occupying SpecDP and licensing a null definite D via agreement), see (41a), but, as quantitatives may always license a null indefinite D, one cannot rule the analysis in (41)b for DP-initial *cei mai mulți* ‘SUP more many’ (in (41)b, the projection dedicated to quantity modifiers is labeled MeasP, following Solt 2009).

- (41) a. [DP [DegP *cel mai* AP] [D  $\emptyset$  [t<sub>DegP</sub> NP]]]  
 b. [DP [D $\emptyset$ <sub>indef</sub>] [MeasP [*cei mai mulți*] [[Meas $\emptyset$ ] NP ]]]

Romanian not only allows *cei mai mulți* ‘most’ in the immigrants-example, it also allows upstairs de dicto readings in non-c-commanded positions, as illustrated in (42), which has been found acceptable by 10 informants out of 11:<sup>x</sup>

- (42) **Cele mai multe pagini** trebuie scrise la literatură.  
 SUP COMP many pages must written at literature  
 ‘The largest number of pages must be written for the literature (course).’

Note moreover that even examples with non c-commanded *the fewest*, deemed impossible by Farkas & Kiss (2000), can be produced in Romanian:

- (43) **Cei mai mulți colegi** au votat pentru Andrei, și **cei**  
 SUP COMP many colleagues have voted for Andrei and SUP  
**mai puțini** pentru Ion.  
 COMP few for Ion  
 ‘The largest number of colleagues voted for Andrei, and the fewest for Ion.’

However, NP-internal correlates, as in (34) above, are totally impossible in Romanian ((44) cannot have an interpretation where the number of chocolate cookies (s)he ate is compared to the number of other sorts of cookies that (s)he ate).

- (44) # A mâncat cele mai multe [prăjituri de ciocolată].  
has eaten SUP COMP many cakes of chocolate

In sum, relative superlatives may vary along three parameters: c-command requirement, allowance of NP-internal correlates, use of a definite determiner. The allowance of upstairs *de dicto* reading might be a fourth parameter. An adequate theory of relative superlatives should check for possible correlations between all these parameters – a very complex task, which goes beyond the scope of this article. From the data presented here, a first important conclusion is that *quantity superlatives tend to occur as indefinite*, even in languages where the article is required for quality relative superlatives – among the languages discussed here, this is the case of Romance languages, English, Albanian.<sup>xi</sup> The existence of languages with the split in (45) is particularly telling for the theory of relative superlatives:

- (45) quality relative superlatives => definite  
quantity relative superlatives => indefinite

This split clearly indicates that the internal structure of the DP is relevant for the phenomenon of -EST raising: if -EST occurs in SpecMeasP (the position occupied by quantity modifiers) it can raise without problems (hence the formally indefinite quantity superlatives). If it occurs lower down, in these languages D must be definite. Depending on how we analyze quality definite relative superlatives, there are two possible explanations (leaving aside Bumford's analysis, which cannot cover indefinite superlatives at all): either (i) -EST raising is actually blocked, and quality relative superlatives are derived via the in-situ strategy, or (ii) raising is performed with the help of a definite D that provides an intermediate step, required for accessing the final DP-external scope position. In this second hypothesis, we may view the article as an expletive whose +def feature agrees with the definiteness feature of the -EST operator, assuming that the degree maximality introduced by the superlative is the counterpart of definiteness in the nominal domain. This expresses the intuition between Krasikova's (2012) analysis, without



making the definite article itself part of the DegP – see (46), which represents the steps of the LF-movement of -EST:

- (46) Foc/Wh [-EST ... [DP [~~EST~~<sub>def</sub>] [D<sub>udef</sub> [...[DegP ~~EST~~ AP]...N]]]]  
John [-EST [ read [DP [~~EST~~<sub>def</sub>] [D<sub>udef</sub> [NP[DegP best] novel]]]]]]

For quantity superlatives, this step is not needed because the quantity modifier, being already a functional item in the periphery of the noun phrase, is closer to the target of -EST raising; therefore, the intermediate step in (46) is not needed for moving -EST out of the DP. A precise implementation of this idea depends on the analysis of the functional structure of DPs with quantity expressions. We may assume that -EST occurs at the edge of the DP in overt syntax already – either the null indefinite D attracts the quantitative to its Spec (see (47)), or, alternatively, (weak) indefinites headed by quantity modifiers have no D layer above the MeasP layer (this latter option is natural in a property-analysis of bare nouns and weak indefinites, see van Geenhoven 1996, McNally 1998, Dobrovie-Sorin & Beyssade 2004, a.o.), see (48):

- (47) Foc/Wh [-EST ... [DP [DegP mo-~~EST~~] [[DØ] [MeasP t<sub>DegP</sub> NP]]]]  
John [-EST [has [DP mo-~~EST~~ [ [DØ] [MeasP ~~most~~ friends]]]]]]

- (48) Foc/Wh [-EST ... [MeasP [DegP mo-~~EST~~ NP]]]  
John [-EST [has [MeasP mo-~~EST~~ friends]]]]

Supporting evidence for the hypothesis (ii), according to which -EST raises through the Spec of definite superlatives, comes from the observation, due to Schwarz (2005), that relative readings are blocked by a DP-initial possessor (see also Chacon & Wellwood 2012, Bumford 2017):

- (49) a. He read my longest paper  
 ≠ He read a paper of mine longer than the papers of mine  
 the others read (Schwarz 2005: ex.41)
- b. # the student who read Shakespeare's longest play  
 (≠ the student who read a play by Shakespeare longer than  
 the plays by Shakespeare the others read)  
 (Bumford 2017: ex.14a)

## 5. The non-c-commanded relative superlatives in the languages of type B

An important consequence of the account presented in the previous section is that, in the languages of type B, the constructions used to render the non-c-commanded relative superlative in (4) (the immigrants-example) do not involve -EST raising.

Recall from section 2 that we have found the following ways of rendering a relative superlative of quantity in a non-c-commanded position, in the example of the form ‘SUP-expression immigrants come from India’, with the intended reading ‘more immigrants come from India than from any other country’:

- (i) THE MOST (Armenian, Basque, Norwegian, Swedish) (ex. (7)b, (8)b, (11)b, (13))
- (ii) Expressions of the type THE LARGEST NUMBER/QUANTITY (English, Spanish) (ex. (16)b, (22)d)
- (iii) THE LARGEST PART (French, Italian, Catalan) (ex. (22)a-c)
- (iv) THE MAJORITY (Spanish, Portuguese, Albanian) (ex. (22)d’-f)

Note that in all these examples the superlative DP is definite, as expected in the case of an in-situ strategy of deriving the relative reading. The possibility that relative readings are derived by both a raising and an in-situ strategy is not rejected in Heim (1999) and is argued for in Pancheva & Tomaszewicz (2012). A type of data discussed in Heim (1999) and Sharvit & Stateva (2002), which suggests that both strategies are involved in quality relative superlatives, is the evaluation of the example (48) in the indicated context:

- (50) JOHN climbed the highest mountain:  
Context: John climbed 2 equally high mountains, which are higher than those climbed by the others

The in-situ analysis predicts that this sentence should be undefined, as there is no unique highest mountain climbed by John. The raising analysis predicts this sentence to be judged as true, without any problem. Apparently, speakers go both ways in the interpretation of this sentence, suggesting that they may employ one strategy or the other.<sup>xii</sup>

The in-situ strategy in the immigrants-example is facilitated by the fact that the referent of the superlative DP is included in a given entity (the total sum of immigrants in the country) and the association with focus

defines a partition of this entity, whose cells correspond to the total number of immigrants from each country. Note that a partition-based strategy has also been proposed for proportional (majority) readings (see Hoeksema 1983, Coppock 2019, Dobrovie-Sorin & Giurgea 2021): the majority reading can be obtained based on a superlative inside a definite phrase if the comparison class of the superlative is set to a binary partition of the total sum of N (or of the entity in which the external argument is included, see the partitive constructions of the form *the most/majority of DP*), and this binary partition is introduced as a variable that is bound by existential closure outside the DP (see Dobrovie-Sorin & Giurgea 2021 for details).<sup>xiii</sup> In the immigrants-example, the partition is not unspecified, as in majority readings, but is provided by association with focus.

Expectedly, the expressions used to render the “immigrants” example are often expressions used as majority quantifiers – this is the case of THE MOST in Basque, Norwegian, Swedish, and THE LARGEST PART /THE MAJORITY in French, Italian, Ibero-Romance, and Albanian. As argued in Dobrovie-Sorin & Giurgea (2021), nouns of the type MAJORITY can be decomposed into a superlative component and a PART-component (insofar as they allow genuine superlative readings).

For concreteness, for in-situ relative superlatives we may adopt the in-situ strategy suggested by Heim (1999), see (32) above (here, we must add the assumption that *maximal* sums of entities satisfying the predicate are compared, which might be a general property of quantity superlatives).<sup>xiv</sup>

- (51) [the -EST-C [d-many immigrants]] [ $\lambda x$  [ $\sim S$  [x come from [India]<sub>F</sub>]]]  
 $C = \cup S \subseteq \{x: \exists y. x \text{ comes from } y\}$   
 $C = \{x: \exists y [\text{country}(y) \wedge x = \sigma z. z \text{ comes from } y]\}$

A further question that arises here is why only some languages use the in-situ strategy with the superlative of MANY/MUCH (type B1 languages). Why is this strategy unavailable in type B2 languages (which, as we have seen, resort to expressions of the type THE LARGEST NUMBER/QUANTITY/PART)?

In order to find a solution, we should have a look at the general properties of superlatives in these languages. We immediately observe that all the languages of type B2, except English, lack dedicated superlative morphology, using instead comparatives embedded in definite DPs. The use of a definite article form as a superlative marker, which has been generalized in Romanian, is restricted in these languages: as we have seen in section 2, in French this use only occurs in predicative and adverbial

positions (and also in postnominal superlatives, but these can be analyzed as predicates of reduced relatives, see Loccioni 2018), but not in prenominal positions – as shown in (52), a cardinal or ordinal may intervene between the article and MORE+AP constituent, showing that the article is not part of the superlative.<sup>xv</sup>

- (52) a. les deux plus hautes montagnes  
 the two more high mountains  
 ‘the two highest mountains’  
 b. la deuxième plus haute montagne  
 the second more high mountain  
 ‘the second highest mountain’

In Italian and Spanish, the only superlative-like use of the article appears in constructions of the type *the X-est (N) possible*, which arguably involve a maximalizing degree relative clause, see Loccioni (2018).

Given the absence of distinct superlative morphology, it is reasonable to assume that the syntactic environment is relevant for licensing a superlative interpretation of a comparative adjective. In particular, prenominal superlatives may be assumed to be licensed in a dedicated projection SupP, as proposed by Loccioni (2018). Evidence for a special prenominal position of superlatives comes from differences in interpretation between superlatives and non-superlative prenominal adjectives. For instance, whereas for most quality adjectives the prenominal position is associated with a non-restrictive interpretation, with superlatives this special interpretation is suspended:<sup>xvi</sup>

- (53) a. l’intéressant roman (marked, non-restrictive) (Fr.)  
 the interesting novel  
 b. le plus intéressant roman (restrictive)  
 the more interesting novel  
 ‘the most interesting novel’
- (54) a. le notevoli palazzi (marked, non-restrictive) (It.)  
 the noteworthy palaces  
 b. le più notevoli palazzi (restrictive)  
 the more noteworthy palaces  
 ‘the most noteworthy palaces’

Moreover, the specifier of this dedicated projection SupP appears to be a scope position, triggering an absolute reading of the superlative: thus,

according to Cinque (2010), the prenominal superlative in (55) only has the absolute reading; for the relative reading, a postnominal superlative must be used.

- (55) Chi ha scalato la più alta montagna innevata?  
 who has climbed the more high mountain snowy  
 ‘Who climbed the highest snowy mountain?’  
 ✓ absolute, \* relative

(Cinque 2010, ch. 2 ex. 23)

Under this hypothesis, the absence of an in-situ strategy for quantity superlatives can be explained by the fact that the quantity adjective is base-generated above SupP, so it cannot move to SpecSupP, the dedicated position of prenominal superlatives:

- (56) [D [<sub>MeasP</sub> MORE [<sub>Meas</sub><sup>0</sup> [<sub>SupP</sub> Sup [.. NP]]]]]

Evidence for the sequence of functional categories D–Meas–Sup comes from the order between superlatives and cardinals shown in (52)a above. The reverse order is not acceptable:

- (57) \* les plus hautes deux montagnes (Fr.)  
 the more high two mountains

## 6. Conclusions

This paper contributes several things to the ongoing debate on the analysis of relative superlatives: (i) indefinite relative superlatives exist even in languages where, based on quality superlatives, it was believed that relative superlatives are only definite; (ii) the fact that relative superlatives allow more easily the absence of the definite article in the case of quantitatives, as compared to quality adjectives, indicate the existence of syntactic constraints on -EST raising out of the DP; I suggested that the definite article in the case of quality relative superlatives is used in order to provide an intermediate position for -EST raising and does not have interpretable definiteness, but surfaces as definite due to agreement with the -EST operator in its Spec; quantity superlatives are base-generated in a more accessible position than quality superlatives, therefore this intermediate step is not needed; (iii) in certain languages, but not in all, -EST raising out of the DP requires that the correlate c-command the superlative in surface structure; (iv) relative

readings of superlatives can be obtained both via -EST raising and via an in-situ strategy. This appears very clearly in languages where there are syntactic constraints on indefinite superlatives, which indicate constraints on -EST raising: the relative superlatives that do not comply with these constraints cannot involve raising; they can only be derived via the in-situ strategy. Expectedly, they are always definite.

## References

- Artiagoitia, Xabier. 2002. "The functional structure of the Basque Noun Phrase." In *Erramu Boneta: Festschrift for Rudolf P. G. de Rijk*, edited by X. Artiagoitia, P. Goenaga & J.A. Lakarra, 73-90. Bilbao: Supplements of ASJU.
- Bumford, Dylan. 2017. "Split-scope definites: Relative superlatives and Haddock descriptions". *Linguistics and Philosophy* 40 (6): 549-593.
- Chacón, Dustin and Alexis Wellwood. 2012. "A superlative puzzle for Bošković's NP/DP parameter". Paper presented at the *Workshop on Languages with and without Articles* on 3.11.2011, handout available at <http://archive.sfl.cnrs.fr/Journees-LSALAA-2012-Langues-avec,1062.html>.
- Cinque, Guglielmo. 2010. *The syntax of adjectives: a comparative study*. Cambridge, MA: MIT Press.
- Cornilescu, Alexandra. 2007. "Double definite constructions in Romanian as evidence for an internal nP phase". Ms., University of Bucharest.
- Cornilescu, Alexandra and Ion Giurgea. 2013. "The adjective". In *A Reference Grammar of Romanian. I: The Noun Phrase*, edited by Carmen Dobrovie-Sorin, and Ion Giurgea, 355-529. Amsterdam, Philadelphia: John Benjamins.
- Coppock, Elizabeth. 2019. "Quantity Superlatives in Germanic, or 'Life on the Fault Line Between Adjective and Determiner'". *Journal of Germanic Linguistics* 31 (2): 109-200.
- Coppock, Elizabeth & David Beaver. 2014. "A superlative argument for a minimal theory of definiteness". In *Proceedings of Semantics and Linguistic Theory (SALT) 24*, edited by Todd Snider, Sarah D'Antonio and Mia Weigand, 177-196. Washington: LSA and CLC Publications.
- Coppock, Elizabeth & Christian Josefson. 2015. "Completely bare Swedish superlatives". In *Proceedings of Sinn und Bedeutung 19*, edited by Eva Csipak & Hedde Zeijlstra, 179-196. Göttingen: University of Göttingen.

- Coppock, Elizabeth, Golsa Nouri-Hosseini, Elizabeth Bogal-Allbritten and Saskia Stiefeling. 2017. "Proportional implies relative: A typological universal". In *Proceedings of the Linguistic Society of America (LSA)*, edited by Patrick Farrell, 2 (12) :1-15.
- Croitor, Blanca & Ion Giurgea. 2016. "Relative superlatives and Deg-raising". *Acta Linguistica Hungarica* 63 (4): 1-32.
- Dobrovie-Sorin, Carmen. 2013. "Most: The view from mass". In *Proceedings of the 19<sup>th</sup> Amsterdam Colloquium*, edited by Maria Aloni, Michael Franke and Floris Roelofsen, 99-107. Amsterdam: ILLC.
- Dobrovie-Sorin, Carmen. 2015. "Two types of most". *Proceedings of Semantics and Linguistic Theory (SALT)* 25: 394-412.
- Dobrovie-Sorin, Carmen & Claire Beyssade. 2004. *Définir les indéfinis*. Paris: CNRS Editions.
- Dobrovie-Sorin, Carmen and Ion Giurgea. 2021. *Majority Quantification and Quantity Superlatives. A Crosslinguistic Analysis of Most*. Oxford: Oxford University Press.
- Farkas, Donka and Katalin É. Kiss. 2000. "On the comparative and absolute readings of superlatives". *Natural Language and Linguistic Theory* 18: 417-455.
- van Geenhoven, V. 1996. *Semantic Incorporation and Indefinite Descriptions: Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. Ph.D. diss., University of Tübingen. Published in 1998 by CSLI.
- Giurgea, Ion. 2013a. *Originea articolului posesiv-genitival al și evoluția sistemului demonstrativelor în română*. Bucharest: Editura Muzeului Național al Literaturii Române.
- Giurgea, Ion. 2013b. "The Syntax of Determiners and Other Functional Categories". In *A Reference Grammar of Romanian. I: The Noun Phrase*, edited by Carmen Dobrovie-Sorin & Ion Giurgea, 99-108. Amsterdam, Philadelphia: John Benjamins.
- Grevisse, Maurice. 2008. *Le bon usage*. Bruxelles: De Boeck; [Louvain-la-Neuve] Duculot.
- Hackl, Martin. 2009. "On the grammar and processing of proportional quantifiers: most versus more than half". *Natural Language Semantics* 17 (1): 63-98.
- Heim, Irene. 1985. "Notes on comparatives and related matters". Unpublished manuscript, University of Texas, available at Austin. <http://semanticsarchive.net/Archive/zc0ZjY0M>.
- Heim, Irene. 1999. "Notes on Superlatives". Unpublished manuscript, MIT, downloadable from

- <http://semanticsarchive.net/Archive/TI1MTlhZ/>
- Hoeksema, Jack. 1983. "Superlatieven". *TABU* 13: 101-106.
- Kotek, Hadas, Edwin Howard, Yasutada Sudo, and Martin Hackl. 2011. "Three Readings of *Most*". *Proceedings of SALT 21*: 353-372.
- Krasikova, Sveta. 2012. "Definiteness in Superlatives". In *Logic, Language and Meaning. 18th Amsterdam Colloquium. Amsterdam, The Netherlands, December 19-21, 2011, Revised Selected Papers*, edited by M. Aloni, V. Kimmelman, F. Roelofsen, G. Sassoon, K. Schulz, and M. Westera, 411-420. Berlin, Heidelberg: Springer.
- Loccioni, Nicoletta. 2018. *Getting "the most" out of Romance*. PhD diss., University of California.
- McNally, Louise. 1998. "Existential Sentences without Existential Quantification". *Linguistics and Philosophy* 21 (4): 353-392.
- Pancheva, Roumyana and Barbara Tomaszewicz. 2012. "Cross-linguistic Differences in Superlative Movement out of Nominal Phrases". In *Proceedings of the 30st West Coast Conference on Formal Linguistics*, edited by N. Arnett and R. Bennett, 292-302. Somerville, MA: Cascadia Proceedings Project.
- Sharvit, Yael and Penka Stateva. 2002. "Superlative expressions, context, and focus". *Linguistics and Philosophy* 25 (4): 453-505.
- Schwarz, Bernhard. 2005. "Modal superlatives". *Proceedings of Semantics and Linguistic Theory (SALT)* 15:187-204.
- Schwarzschild, Roger. 2006. "The role of dimensions in the syntax of noun phrases". *Syntax* 9 (1): 67-110.
- Solt, Stephanie. 2009. *The semantics of adjectives of quantity*. PhD diss., City University of New York.
- Szabolcsi, Anna. 1986. "Comparative superlatives". *MIT Working Papers in Linguistics* 8: 245-66.
- Teodorescu, Viorica Alexandra. 2009. "Modification in the Noun Phrase. The Syntax, Semantics, and Pragmatics of Adjectives and Superlatives". PhD Diss., The University of Texas at Austin.
- Wilson, E. Cameron. 2018. "Amount Superlatives and Measure Phrases". PhD Diss., The City University of New York.
- Živanović, Sašo. 2007. "Quantificational Aspects of LF". PhD diss., University of Ljubljana.

## Notes

---

<sup>i</sup> The initial wording was "absolute vs. comparative", used in Szabolcsi (1986). The term "relative" appears two times in Heim (1999) and has been adopted by much of the more recent literature (e.g. Hackl 2009, Pancheva & Tomaszewicz



2012, Coppock & Beavers 2014, Bumford 2017). I prefer this term because comparison is always involved in superlative readings, including the absolute one, and the opposite of “absolute” is “relative”.

<sup>ii</sup> The amendment consists in the fact that, when assessing that other entities in the comparison class do not have the property to the degree to which the comparandum has it, difference (non-identity) is replaced with non-overlap, i.e., instead of ‘x has P to the degree d and for all y in C such that  $y \neq x$ , y doesn’t have P to the degree d’, the formula would say that ‘x has P to the degree d and for all y in C such that y *does not overlap with* x, y doesn’t have P to the degree d’. Hackl (2009) shows that with this amendment, the absolute reading of *most* amounts to the proportional reading. Given that the proportional reading of quantity superlatives is much rarer crosslinguistically than the relative reading, I conclude that Hackl’s amendment of the general semantics of superlatives should be discarded.

<sup>iii</sup> The questionnaires consisted in a number of sentences to be translated from English, comprising examples with comparative *more*, superlative *most* and proportional *most* in several environments. The number of informants for each language varied from 1 to 6. We tested 28 languages.

<sup>iv</sup> Data provided by Alexander Grosu and Omer Preminger.

<sup>v</sup> With quality superlatives, the definite article is optional, rather than ruled out, as in the case of relative superlative *flest* ‘most’ in (10)b:

- (i) Gloria sålde godast glass. (Coppock & Josefson 2015: ex. 28)  
 Gloria sold best ice.cream  
 ‘Gloria sold the best icecream’: relative only
- (ii) Fredrik köpte det dyraste vin-et. (ibid.: ex. 29)  
 Fredrik bought the expensive-sup wine-the  
 ‘Fredrik bought the most expensive wine’: relative or absolute

<sup>vi</sup> The reason for this oscillation will be clear when the analysis is presented. Anticipating, I will argue that the article-less version relies on -EST raising and the version with the article is based on an in-situ scope of -EST. If the language does not impose an overt c-command requirement on -EST raising, both types of superlatives may occur in the non-c-commanded position. The use of the article for the proportional reading blocks the superlative reading for some speakers.

<sup>vii</sup> For more on this type of predicative superlatives, see Croitor & Giurgea (2016). For other French examples of the type in (17), see Loccioni (2018:96-97).

<sup>viii</sup> The same use of THE LARGEST PART for (4) was provided by our Breton informant (but this can be due to French influence).

<sup>ix</sup> The only fully-fledged account I know of is Sharvit & Stateva (2002), according to whom the superlative DP in upstairs de dicto contexts such as (25)c has a very peculiar semantics, denoting a property obtained by applying two operators, IDENT-w and a property-level iota, to the property ‘be a d-high mountain’; as a result, from a contextually set of properties containing degrees (e.g. *be a 5000ft mountain, be a 4000ft mountain, be a 3000ft mountain*), the one is selected that has the same extension, in the worlds considered, as the property *be the highest*

*mountain*. The predicate (e.g. *climb*) combines with this property via type-shifting, whereby the external argument of the object is existentially bound.

<sup>x</sup> In English, this type of sentence, with *the most*, is expected to be deviant because of the ban on non-c-commanded *the most N*. However, the sentence of this type I tested, given in (i), was not rejected by all speakers: it was judged fine by 2 informants, somehow not perfect by 3 informants, very strange but not ungrammatical by one informant, and ungrammatical by one informant.

(i) %The most books must be read for the HISTORY course

<sup>xi</sup> In Swedish, where for quality relative superlatives definiteness marking is optional, for quantity relative superlatives, at least in the c-commanded position, definiteness marking is ruled out, according to Coppock's (2019) investigation.

<sup>xii</sup> Sharvit & Stateva (2002) actually take this fact as evidence in favor of the in-situ analysis, claiming that those speakers who find (50) true simply ignore one of the mountains climbed when they establish C, but I cannot see why they should ignore it. It seems more likely, to me, that they use the raising strategy, directly comparing individuals in terms of the height of climbed mountains.

<sup>xiii</sup> The fact that majority readings are restricted to some languages and lexical items (e.g. *largest part*, but not *most*, in some languages) is explained in Dobrovie-Sorin & Giurgea (forth.) by making the choice of an unspecified binary partition as comparison class a matter of lexical specification. This contrasts with Hackl's (2009) analysis, which predicts that *any* absolute superlative of *many/most* should yield the majority reading.

<sup>xiv</sup> A similar analysis has been proposed for the superlative use of bare *most* in non-c-commanded positions in examples of the type in (16)c by Kotek et al. (2011). For a more detailed discussion of this study (which was unknown to me when I wrote the first draft of this article and the presentation on which this article is based), see Dobrovie-Sorin & Giurgea (2021), chapter 4 section 8.

<sup>xv</sup> Compare Romanian, where we never find a comparative without *cel* in a superlative function – either the whole constituent occurs before the numeral (this is current with cardinals) or the constituent [*cel* + comparative] appears after the numeral (this is normal with ordinals):

- (i) [cei mai înalți] doi munți / \*cei doi mai înalți munți  
 SUP COMP high two mountains the two COMP high mountains
- (ii) al doilea [cel mai înalt] munte / \*al doilea [mai înalt] munte  
 the second SUP COMP high mountain the second COMP high mountain

<sup>xvi</sup> See Loccioni (2018:chapter 2) for other contrasts of this type.

## CHAPTER NINE

# ROMANIAN *CÂTE* AS A NOMINAL AND ADVERBIAL MARKER OF DISTRIBUTIVITY

MARA PANAITESCU

The paper focuses on the syntactic and semantic properties of markers of nominal distributivity, with particular reference to Romanian marker *câte*. It will be shown that the two syntactic positions analyzed (adverbial and nominal) correlate with two different semantic construals. These two construals have been argued for in the investigation of various distributive markers in a wide variety of unrelated languages, in certain cases with different authors disagreeing on the choice between the two when looking at one and the same marker in one and the same syntactic construction. The contribution of this work is to establish a link between syntactic position and distributional constraints which require a minimum of item-specific stipulations.

Keywords: distributive marker, distributive operator, share, key, partition, pluractionality

### 1. Introduction

The semantic status of nominal distributive markers on the share is still a subject of debate in the literature on plurality. This may be due to the fact that various markers in various languages have been the target of a unitary investigation, as well as the fact that other connected semantic and pragmatic phenomena may interact with these markers in producing a certain observable effect. The primary goal of this paper is a modest one. Looking at Romanian distributive marker *câte*, I will try to show that the two syntactic positions that are available to it translate as two different semantic contributions for this item. In its guise as a VP adjunct, *câte* is in the immediate scope of a partition operator proper. This operator ranges

over a plurality of events described by the main predicate as well as pluralities of entities. The role of the distributive marker is to specify the size (cardinality) of the cells. In its guise as a DP constituent, *câte* is a marker of nominal distributivity on the share. This means it merely signals the presence of a (possibly covert) operator which takes the DP containing *câte* in its scope and distributes over it. The secondary goal, which will turn up as the discussion unveils, is to contribute to the discussion surrounding the distributivity of unmarked cardinal expressions (i.e., those cardinal expressions that have no share marker on them) in general. The paper is organized as follows. Section 2 introduces the two constructions: 2.1 presents the distribution of *câte* as a nominal constituent, while 2.2 illustrates the distribution of *câte* as an adverbial. Section 3 puts forth an account of the two constructions 3.1 addresses the adverbial construction as being in the immediate scope of a partition operator, while 3.2 analyses the nominal construction as being in the scope of a distributive operator or a partition operator. Section 4 concludes the paper.

## 2. On distributivity: key markers, share markers and pluractional markers

### 2.1 Nominal distributivity as a key-share relation

This subsection presents some general notions regarding nominal distributivity and introduces the Romanian marker *câte* as a cardinal/measure share marker. It must be pointed out explicitly from the beginning that there is another use of *câte*, namely its occurrence inside singular indefinite DPs. For a contrast between the two, see Panaitescu (2018). The focus of this paper will be only *câte* + Cardinal and *câte* + Measure constructions.

A first remark is that all of the instances of distributivity discussed here are phrasal distributivity. This is not to be confused with lexical distributivity, which is understood as a lexical semantic property of verbs<sup>1</sup>. For instance, a verb like *smile* is lexically distributive because if two people smiled then it can be asserted that each of them did. Conversely, *gather* is not distributive, because if ten people gathered, it cannot be said that each of them did. The term distributivity which will be used throughout is meant to capture phrasal distributivity. For instance, this kind of distributivity may be introduced by quantifier *each*. The effect of *each* on a verb which is lexically underspecified with respect to distributivity is to enforce the distributive reading. The verb *eat* is taken as an illustration.

- (1) The children ate one pizza.
- (2) Each child ate one pizza.

In (1), it may be that one pizza was eaten overall and each child ate a piece or that each child ate one pizza, with a total of eaten pizzas equaling the number of children. In (2), only the latter interpretation is possible, also due to the fact that the verb is non-iterative (the same pizza cannot be eaten multiple times). Thus, it may be said that the quantifier marking the subject enforces a distributive reading on the object.

In a language such as Romanian, the “distributtee” (in this case the direct object) can be marked for distributivity:

- (3) Copiii au mâncat câte două pizze  
children.DEF have eaten DIST two pizzas  
'The children ate two pizzas.'  
(only interpretation available: they ate two pizzas each)
- (4) Fiecare copil a mâncat câte două pizze.  
each child has eaten DIST two pizzas  
'Each child ate two pizzas.'

In the rest of the paper, I will follow the terminology of Choe (1987) and call the phrase which induces distributivity the sortal key and the nominal phrase which co-varies with it the distributive share. Going back to the Romanian share marker, first of all, the marking with *câte* in these examples is optional. In (3), the effect of adding the marker is a disambiguation between the two available readings discussed above for the English counterpart (1). The only available reading for (3) is the distributive one. In (4), however, the reading is unambiguously distributive to begin with (with or without *câte*).

With an iterative verb, marking the subject with *each* enforces a plurality of events but does not exclude a so-called wide scope reading of the cardinal indefinite in direct object position (Braşoveanu & Farkas 2011).

Adding *câte* excludes this wide scope reading, namely the interpretation in which there is a plurality of events involving the same participant over and over again. To take an example:

- (5) Profesorii au corectat zece examene. Sunt acolo, pe masă  
 professors.DEF have corrected ten exams. are there on table  
 toate zece.  
 all ten  
 ‘The professors have corrected ten exams. They are there on the table,  
 all ten of them.’

The sentence is ambiguous between a cumulative reading and a distributive one. On the cumulative reading, the total number of corrected exams is ten and each professor corrected a part. On the distributive reading, given the continuation, each professor took the ten exams and corrected them in turns. This is the wide scope reading of the direct object. With *câte* on the other hand, this reading is excluded. The only interpretation available is that more than ten exams were corrected overall (although incidentally, some of them may have been corrected by two professors).

- (6) Profesorii au corectat câte zece examene.  
 professors.DEF have corrected DIST ten exams.  
 ??Sunt acolo, pe masă toate zece.  
 are there on table all ten  
 ‘The professors have corrected ten exams. They are there on the table,  
 all ten of them.’

The same happens with a quantified subject such as *Each professor corrected ten exams*. This sentence allows for a wide scope interpretation of *ten exams*. Again, the Romanian counterpart with *câte* does not:

- (7) Fiecare profesor a corectat câte zece examene.  
 each professor has corrected DIST ten exams.  
 ‘Each professor has corrected ten exams.’  
 (only interpretation: ten exams per professor)

The key need not be nominal, it can also be temporal, eventive or spatial:

- (8) Am citit (*câte*) trei cărți pe săptămână.  
 have.IP read DIST three book on week  
 ‘I/We read three books per week.’

- (9) De fiecare dată când o vizitează, Ion îi  
 of each time when CL.3P.SG.F.ACC visits Ion CL.3P.SG.DAT  
 recomandă Mariei câte două seriale.  
 recommends Mary. DAT DIST two shows  
 ‘Each time he visits her, John recommends two (different) shows to Mary.’  
 (Only interpretation: Ion recommends two possibly different shows each time)
- (10) În două locuri erau câte trei greșeli grave.  
 in two places were DIST three mistakes serious  
 ‘In two places there were three serious mistakes.’

In (8) above, the temporal ratio expression *pe săptămână*, “per week” serves as key. The contribution of *câte* here is not detectable, as there is no difference in meaning with or without it: different groups of three books were read over one-week periods. In 0, the specific reading is excluded. I take *per week* to range over time intervals and, following Rothstein (1995), *each time* to range over events. In (10) *in two places* ranges over locations. In this case also, there is no disambiguation effect of the distributivity marker.

Also, the key need not be the subject and the share need not be the direct object, the subject-direct object case is just the most at hand for illustrations. In (11) below, the total amount of money is 30 lei, distributed over the three beggars and in (12) the share is a prepositional/ indirect object.

- (11) Am dat câte 10 lei la trei cerșetori  
 have. 1P given DIST 10 lei to three beggars  
 ‘I gave three beggars 10 lei each.’
- (12) Fiecare profesor a vorbit cu câte doi studenți/ a dat 10  
 each professor has talked with DIST two students/ has given 10  
 la câte doi studenți.  
 to DIST two students  
 ‘Each professor talked with/ gave an A to two students.’

It has already been mentioned that *câte* is never obligatory and that it generally disambiguates in favor of a co-variation reading of the nominal phrase it is a part of, but this is not true in all cases. Conversely, not all environments allow the use of *câte*. In a nutshell the DP containing this marker must be in the scope of some plurality or distributive quantifier.

- (13) ??Marcel a mâncat *câte* două pizze.  
 ‘Marcel has eaten DIST two pizzas.’

The sentence above lacks a distributive key and is therefore uninterpretable out of the blue. It is not ungrammatical, as the English sentence *\*He ate two pizzas each* is. This is because the key can always be covert and pragmatic considerations further complicate the picture. For instance, as part of a discussion regarding what Marcel ate each day, (13) becomes perfectly acceptable. These are properties that bring *câte* close to reduplication in Hungarian in Braşoveanu and Farkas (2011), Farkas (2015) and Kaçchikel (Henderson 2014, 2016). What matters for licensing is the existence of some plurality. This plurality may even be covert, as in the example below, which is only interpretable as the pairs of soldiers standing on guard in shifts (a temporal partition).

- (14) *Câte* doi soldaşi vor păzi comoara până mâine.  
 DIST two soldiers will guard treasure. DEF until tomorrow  
 ‘Two soldiers will guard the treasure until tomorrow.’  
 (Interpretation: pairs of soldiers will guard the treasure until tomorrow)

To anticipate, the acceptability of *câte* will be explained by the presence of a contextually supplied partition as defined in Champollion (2016). As for the syntactic position of *câte*, I will assume it is inside the NumP projection (in the general case inside a Quantity/ Measure Phrase).

## 2.2 *Câte* as a pluractional adverbial

This subsection discusses *câte* in its guise as a pluractional marker<sup>2</sup>, as found within pluractional adverbial phrases. Used adverbially, *câte* can target a plural argument of the verb. The semantic effect is to partition the main event into subevents in which members of the targeted plurality are participants.

- (15) Nuntaşii dansau *câte* trei.  
 wedding-guests. DEF danced DIST three  
 ‘The wedding guests were dancing in threes.’
- (16) Cei doisprezece copii s-au grupat *câte* trei.  
 DEF twelve children REFL-have grouped DIST three  
 ‘The twelve children made groups of three.’



In (15) above, the interpretation is the following: the maximal plurality of wedding guests is involved in a global event and this event can be divided into subevents which are events of dancing and which each contain three members of the aforementioned plurality of guests. Here, incidentally, the totality of wedding guests also happens to be the Agent of a dancing event. This is because of the lexical properties of the verb *dance*. The same mechanism of interpretation applies to (16) though, where we can see that, due to the verb, the global event is built as a sum of grouping events (a plurality of events) such that it cannot be said that the total plurality of children are the Agent of a grouping event. Rather, they are the Agents of a plurality of grouping events.

If there are more than two nominal constituents, the pluractional adverb can target either of the two:

- (17) a. Copiii mănâncă pizzele câte doi.  
 children.DEF.PL.M eat pizzas. DEF.PL.FEM DIST two. PL.M
- b. Copiii mănâncă pizzele  
 children.DEF.PL.M eat pizzas. DEF.PL.FEM  
 câte două.  
 DIST two. PL.FEM  
 ‘The children eat pizzas in twos.’

In (a) above, the adverbial agrees in gender with the subject, so it reads that the children shared pizzas in pairs. In (b), each child eats two pizzas.

There is also a reduplicated construction, translated with the English *one by one, two by two* etc.

- (18) Spectatorii ies unul câte unul/ doi câte doi din sală.  
 spectators. DEF exit one DIST one two DIST two from hall  
 ‘The spectators are exiting one by one/ two by two from the hall.’

The only difference from the non-reduplicated construction is that the latter merely allows, while the former requires that there should be a temporal sequence of events. This construction will be left for further analysis and will not be included in the present discussion any further.

The distribution of adverbial *câte* is quite restricted. Similarly to the English construction which was used for the translations (*in twos, in threes* etc.), the sentences that allow this constructions are about some form of organized event and the pluractional adverbial phrase describes how the subevents are organized. Trying to use adverbial *câte* in a situation in which,

say, the children were crying in threes yields an odd result without further contextual support.

An important observation is that, as opposed to nominal *câte*, which relates a plurality of key entities, times or locations to a plurality of share entities via a plurality of events, adverbial *câte* relates a plurality of key entities to a plurality of events and the share is a subplurality of key entities of the specified cardinality.

A second important observation, an apparently trivial one, is that the key cannot be overtly marked with a distributive quantifier.

- (19) \*Fiecare spectator a ieșit *câte* unul.  
 each spectator has exited DIST one

The following section starts out with a proposal for the adverbial uses of *câte*. More specifically, I will propose a cover-based analysis of adverbial *câte*, following Beck and von Stechow (2007). The main idea behind it is that the distributive marker requires to be locally bound by a partition operator. This explains its incompatibility with other distributors such as *each*. The second part of the next section is concerned with the analysis of nominal *câte*.

### 3. Two accounts

#### 3.1 Pluractional adverbial *câte* and partitions

This construction requires the presence of a local plural nominal participant and a partition over the event described by the main predicate into subevents that contain parts of that nominal participant of the specified cardinality. The notion of a partition will be detailed further here, following Beck & von Stechow (2007).

A partition is defined as a cover with no overlapping cells. The following definitions and further implementation onto pluractional adverbials encroached in an event semantics with the types individual  $e$  and eventuality  $v$ , with both domains having a mereological structure.

- (20) a.  $x$  and  $y$  overlap iff they have some common part:  
 $x \circ y$  iff  $\exists z[z \leq x \ \& \ z \leq y]$   
 b.  $x$  and  $y$  are distinct iff they do not overlap

Link's (1983) \* operator for the pluralization of  $\langle e, t \rangle$  predicates is extended to operators pluralizing relations. For two-place relations of event-

entity pairs, the notation is \*\*. Also, the concatenation of an object of type  $e$  with one of type  $v$  forming such pairs is represented between curly brackets as  $\{x, e\}$ ,  $\{y, e'\}$  etc., with  $x, y$  variables of type  $e$  and  $e, e'$  variables of type  $v$ :

(21) Cumulation operator \*\*

Let  $R$  be a relation of type  $\langle e, \langle v, t \rangle \rangle$ . Then  $[**R]$  is the smallest relation  $R'$  such that the conditions in (a) and (b) are satisfied.

a.  $R' \supseteq R$

b. for all  $\{x, e\}$  and  $\{y, e'\}$ : If  $\{x, e\} \in R'$  and  $\{y, e'\} \in R'$ , then  $\{x + y, e + e'\} \in R'$

The definition above says that a cumulation operator takes a relation between an entity and an event and forms a pluralized relation out of it: if a pair formed of an entity  $x$  and an event  $e$  are in the denotation of this relation, and a pair formed of an entity  $y$  and an event  $e'$  are also in the denotation of the relation, then so is the pair formed of the sum  $x+y$  and  $e+e'$ . To illustrate, let's take a transitive relation such as  $\{\text{eat, some potatoes}\}$ . If John ate some potatoes, then  $\{x, e\}$  (where  $x$  are the potatoes eaten and  $e$  is the corresponding eating event performed by John) is an element of the unpluralized relation  $R$ , but also a member of the pluralized relation  $**R$ .

Conversely, if also Mary ate some potatoes, then  $\{y, e'\}$  (where  $y$  are the potatoes eaten by Mary and  $e'$  is the corresponding eating event) is an element of the unpluralized relation  $R$ , but also a member of  $**R$ . Condition (b) says that the sum of the two pairs ( $\{x + y, e + e'\}$ ) is also a member of  $**R$ .

(22) Let  $R$  be a relation of type  $\langle e, \langle v, t \rangle \rangle$ . Then for any  $x, e$ :

$$[**R](x)(e) = 1 \text{ iff } \forall x' \leq x: \exists e' \leq e: R(x')(e') \ \& \ \forall e' \leq e: \exists x' \leq x: R(x')(e').$$

In words, for any (plural) entity  $x$  and plural event  $e$ , the pair  $\{x, e\}$  is in the pluralized relation  $**R$  if and only if all parts of this (plural) entity are related by the unpluralized  $R$  to parts of the plural event and all parts of the (plural) event are related by  $R$  to some part of the (plural) entity.

The size of the parts is assumed by Beck and von Stechow (2007) following Schwarzschild (1996) to be contextually provided by the universe of discourse in the form of a cover. A cover is defined as follows:

(23) **Cover**

Cov is a cover of  $x$  iff Cov is a set such that  $\Sigma \text{Cov} = x$ .

$\Sigma M$  is the fusion of the elements of a set  $M$  if it has all of them as parts and has no part that is distinct from each of them. A cover of  $x$  is therefore a set of elements that are parts of  $x$  (i.e. that fuse into  $x$ ). The covers which are introduced by the pluractional operators discussed here are also partitions, i.e. they contain only non-overlapping elements<sup>3</sup>.

- (24) a. A cover  $Cov$  is a partition iff for any  $x, y \in Cov$ :  $x$  and  $y$  don't overlap.  
 b. **PART(Cov,x)** := 1 iff  $Cov$  is a partition (and a cover) of  $x$ .  
 c.  $Cov[x] = \{y: y \in Cov \ \& \ y \leq x\}$

The final step is to introduce a VP-level cover over a relation between entities and events:

- (25)  $PL = \lambda Cov. \lambda R_{\langle e, \langle v, t \rangle \rangle}. \lambda x_e. \lambda e_v. PART(Cov, e+x) \ \& \ **[\lambda x'. \lambda e'. Cov(e') \ \& \ Cov(x') \ \& \ R(x')(e')](x)(e)$

The operator  $PL$  applies to a cover, a relation  $R$  between individuals and events, a plural entity  $x$  and a plural event  $e$  and requires the cover to be a partition  $e+x$  and that each pair  $\{e', x'\}$  of subparts in the cover should satisfy the relation  $R$ . Pluractional adverbials are modifiers that further constrain the cover. For the present purposes, the pluractional adverbial will specify the cardinality of  $x$ , thus constraining the size of the cells of the cover based on the number of entities in each cell.

The following is an implementation of the formalism described so far by Beck and von Stechow, also partially pursuing the application by Braşoveanu & Henderson (2009) to *one by one* adverbials. The application to pluractional *câte* + Cardinal constructions is the following:

- (26) Nuntaşii                      dansau *câte*    trei.  
 wedding-guests. DEF danced DIST three  
 'The wedding guests danced in threes.'
- a. [the wedding guests] [ $PLCov$  [*câte* three [ $\lambda 2$  [ $t2$  danced]]]]  
 b.  $\exists e \ \& \ Part(Cov, e + THE\_WEDDING\_GUESTS) \ \& \ \{e, THE\_WEDDING\_GUESTS\} \in **[\lambda y. \lambda e. Cov(y) \ \& \ Cov(e) \ \& \ |y| = 3 \ \& \ DANCE(e, y)]$

- (27) Învăţătoarea a grupat elevii                      *câte*    trei.  
 teacher.DEF has grouped students. DEF DIST three  
 'The teacher grouped the students in threes.'

- a. [the students] [*PLCov* [câte three [ $\lambda 2$  [we grouped *t2*]]]]  
 b.  $\exists e \& \text{Part}(\text{Cov}, e + \text{THE\_STUDENTS}) \&$   
 $\{e, \text{THE\_STUDENTS}\} \in^{**} [\lambda y. \lambda e. \text{Cov}(y) \& \text{Cov}(e) \& |y| = 3$   
 $\& \text{GROUP}(e, \text{TEACHER}, y)]$

I will only discuss example (27) but everything carries on to the example (26), where the subject is the plurality whose parts are the cells of the partition. The key entity scopes out of the VP. The plural cover selects the relation formed of the trace of the extracted nominal constituent and the event variable. This relation is further modified by the pluractional adverbial phrase, which specifies the cardinality of the entity variable in the partition. The fact that the partition is contextually provided explains the limited distribution of these constructions. The semantic constraint of *câte* is that it should be distributed over by the pluralization operator taking scope immediately above it.

The following section presents a similar account of nominal *câte*.

### 3.2 Nominal *câte*, D and Part

As shown in section 2.1, nominal *câte* is less constrained in terms of possible licensing possibilities than adverbial *câte*. Nominal *câte* can be licensed under the scope of a plural nominal constituent (28), of a nominal constituent introduced by a distributive universal quantifier (29), of some temporal, eventive or locative adverbial phrase (30), or of some possibly covert partition.

- (28) Copiii au mâncat câte două pizze  
 children.DEF have eaten DIST two pizzas  
 ‘The children ate two pizzas.’  
 (only interpretation available: they ate two pizzas each)
- (29) Fiecare profesor a corectat câte zece examene.  
 each professor has corrected DIST ten exams.  
 ‘Each professor has corrected ten exams.’  
 (only interpretation: ten exams per professor)
- (30) Am citit (câte) trei cărți pe săptămână.  
 have.1P read DIST three books on week  
 ‘I/We read three books per week.’

- (31) *Câte doi soldați vor păzi comoara până mâine.*  
 DIST two soldiers will guard treasure. DEF until tomorrow  
 ‘Two soldiers will guard the treasure until tomorrow.’  
 (Interpretation: pairs of soldiers will guard the treasure until tomorrow)

A second difference is that the share cardinality introduced by the nominal marker applies to the local share NP and not to a key nominal plurality, as is the case for the adverbial construction. For the adverbial construction, the share is a subplurality of the key<sup>4</sup>. This difference is illustrated with the LFs below:

- (32) [**the wedding guests**] [*PLCov* [**câte three** [ $\lambda 2$  [*t2* danced]]]]  
**KEY** **KEY-PART (=SHARE)**
- (33) [**each child**] [*ate* [**câte two pizzas**]]]  
**KEY** **SHARE**
- (34) [*PLCov*-shifts [**câte two soldiers** [*stand-on-guard in shift*]]]  
**KEY-SUM** **SHARE**

In the case of nominal *câte*, the distributivity may be supplied by a universal distributive quantifier. Following Champollion (2016), distributivity comes in two guises, one provided by the pragmatics (partitions) and the other provided in the semantics of sentences with plural constituents. The latter is represented by the D operator, as first introduced by Link 1987. The distinction between the two is due to the granularity of the share. The default granularity is ATOM and it is available in the nominal domain but not in the temporal one. The universal distributive quantifier *each* has incorporated in its semantics a D operator which is comparable to a cover which always sets its cells to the atomic level of granularity. In the case of sentences containing plural definite subjects, there is presumably an optional covert version of D, which explains the ambiguity between the distributive and the cumulative readings discussed in section 2.1.

- (35) The boys ate two pizzas.  
 a. CUMULATIVE: the boys ate a total of two pizzas  
 (i) The boys ate two pizzas.  
 (ii)  $\exists e$  [*\*agent*(*e*) = THE\_BOYS & *\*eat*(*e*) and  
*\*pizza*(*theme*(*e*)) and |*theme*(*e*)| = 2]

- b. DISTRIBUTIVE: the boys each ate two pizzas  
 (i) The boys [ $D_{ag}$  [ate two pizzas]]  
 (ii)  $\exists e$  [ $*agent(e) = \text{THE\_BOYS}$  and  $e \in *\lambda e'(*eat(e')$  and  $*pizza(\text{theme}(e'))$  and  $|\text{theme}(e')| = 2$  and  $\text{atom}(agent(e'))$ ]

In the formalizations above, on the cumulative reading, the pizzas of cardinality two are the theme of the main event of eating, which also contains the boys as agents. On the distributive reading, the distributive operator targets the agent as key. The contribution of  $D_{ag}$  is to specify that the global event  $e$  is identified as a sum of events  $e'$  which satisfy the verbal property. These subevents have as theme two pizzas and an atomic agent. Since the agent of the global event  $e$  is the sum of boys, each subevent  $e'$  can only have one boy as agent. Thus, each boy is in an eating relation to two pizzas.

Moving on to nominal *câte*, it has already been suggested that adverbial *câte* has to be in the immediate scope of a plural cover operator. Thus, by assuming that there always is a covert or overt licenser for nominal *câte* as well, we can offer a unified treatment of this marker as well as explain the numerous cases in which there is an overt distributivity operator which unambiguously targets the share. The difference between (the Romanian translation of) *The boys ate two pizzas* and *The boys ate CÂTE two pizzas* is that in the latter case the D operator is obligatory due to the licensing conditions of the distributive marker. This allows for a unification of the nominal and adverbial uses which comes very close to the proposal in Oh (2001, 2005) for the Korean distributive marker *ssik*, which was analyzed by using an analogy to polarity sensitive items, as a “distributive sensitive item”. The same idea was proposed for nominal *câte* in Panaitescu (2019) and has here been extended to adverbial *câte*.

This account allows for the following generalization. The possible licensers for nominal *câte*, ranging from strongest to weakest, are:

- (36) Distributive operators:  
 (i) D operator  
 (ii) Part operator with explicit key unit  
 (iii) Part operator with contextually provided key

Various positive/ negative polarity licensers have a different capacity to license polarity sensitive items. For instance, negation meets all the requirements to be a licenser for all negative polarity items, while questions have a more limited capacity to license negative polarity items<sup>5</sup>. Similarly, in the case of distributivity inducing operators there seems to be a gradient

in terms of strength when looking at their capacity to license the distributive marker *câte*. Assuming that the D operator is restricted to the nominal domain, the event and temporal domains are only compatible with the Part operator. For instance, in the case of a temporal partition, one needs to set the granularity i.e., establish the size of the cells of the partition. If the way in which to form the cells of the partition is not explicit, sentences containing *câte* sound degraded, as also remarked for Hungarian reduplication in Farkas (2015):

- (37) ??De obicei citesc                      *câte* trei cărți.  
           of custom read. 1P.SG            DIST three books  
           Intended: ‘I usually read three books (and they are different each time).’

In (37) above, the frequency adverb forces the reading event to be interpreted as iterative, but the triplets of books seem to resist co-variation. The same of course can be said about the availability of a non-specific reading of the plain cardinal, as illustrated in the intended translation above: it is very difficult to obtain. This is exactly the conclusion that needs to be stressed: contrary to other accounts which assume that distributive markers bring a contribution to the semantics of the sentences they are a part of, that they differ in meaning from the version with the plain cardinal, or that they induce distributivity themselves<sup>6</sup>, I claim that, at least in the case of the marker studied here, there is no difference between the plain cardinal on a non-specific reading and the marked cardinal. The two go in parallel: if the non-specific reading is unavailable, the distributive marker is infelicitous and vice versa. When the key is a nominal plurality, the operator D is available and thus the distributive marker is used to mark the share.

In the temporal and spatial domain, D cannot apply, so a partition is needed. If the partition is overt (as in the case of (30) above for instance), the marker can be used. If it is covert (as in (31)), the marker strongly depends on world knowledge and contextual information. For spatial key identification, see Knežević & Demirdache (2018), who conduct experiments on Serbian distributive marker *po* involving the visual identification of spatial partitions. Since it has been pointed out along the way that there are several cases in which *câte* does not serve to disambiguate (because a plain cardinal would have been unambiguously non-specific anyway), the natural question to ask is if there is any semantic contribution of this marker. In Panaitescu (2019) I claim that this contribution is an additivity condition. This additivity condition is not “visible” in constructions involving *câte* +



Cardinal, but becomes apparent if the marker combines with other measure phrases.

#### 4. Conclusion

The paper was concerned with Romanian marker *câte* + cardinal in two syntactic positions, one nominal and the other adverbial. The two constructions were shown to involve two different interpretations. The nominal use exhibits the traditionally exemplified form of distributivity associated with, for instance, binominal *each* in the sentence *The boys ate two pizzas each*. The adverbial use, on the other hand, seems to be at the same time collective and distributive. A similar observation was made for English *one by one* in Braşoveanu & Henderson (2009), who call the semantic effect of this adverb “encapsulation”. The sentence containing *one by one* asserts the existence of a big event and introduces all of its global participants, but also specifies that some thematic role as well as the event itself can be divided into smaller events which verify the main predicate. The same seems to apply to adverbial *câte*. Nevertheless, due to examples like (27) with the boys grouping themselves in threes, it can be claimed that the adverbial construction also involves distributivity proper, given that the global event of grouping in this example is necessarily plural (i.e. it cannot be said that all of the boys formed one single group)..

A second remark is that it might seem that the syntactic position of the marker correlates with the interpretation, namely the adverbial creates an impression that there is a global collective event (which is partitioned by the marker in the manner described in section 3.1), while in the case of the nominal marker the interpretation seems to be purely distributive (these sentences assert the existence of a plurality of events). A surprising fact, which at least casts doubt on such a theoretical move, is that according to Hwang (2012), Korean has split and non-split measure phrase constructions (*boy-case 3-cl-ssik* and *boy-3-cl-ssik-case* respectively). The available interpretations are the reverse of what has been seen for Romanian: the non-split constructions seem to allow for collective readings, as in (38), while the split counterparts do not:

- (38) *boy 3-cl-ssik-nom made a chair.*  
 OK: In each subevent, a triplet of boys made a chair. (distributive)  
 OK: In one event, triplets of boys joined forces to make a chair.  
 (collective)

I leave the elucidation of this matter for further study. Another issue which has been addressed only marginally is a full classification of partitions, including the spatial and temporal domain, ranging from frequency adverbs, habituals, pluractional adverbs and so on. This task will also be left for future work.

## References

- Balusu, Rahul. 2006. "Distributive reduplication in Telugu". In *Proceedings of the 36<sup>th</sup> Annual Meeting of the North East Linguistic Society*, edited by C. Davis, A. R. Deal & Y. Zabbal, 39-52. Amherst, MA: GLSA.
- Beck, Sigrid, and Arnim Von Stechow. 2007. "Pluractional adverbials". *Journal of Semantics*, 24(3): 215-254.
- Braşoveanu, Adrian, and Donka Farkas. 2011. "How indefinites choose their scope". *Linguistics and Philosophy* 34 (1): 1–55.
- Braşoveanu, Adrian, and Robert Henderson. 2009. "Varieties of distributivity: One by one vs. each". Accessed March 20, 2021. [http://people.ucsc.edu/~abrsvn/varieties\\_of\\_distributivity\\_SALT\\_19.pdf](http://people.ucsc.edu/~abrsvn/varieties_of_distributivity_SALT_19.pdf).
- Cable, Seth. 2013. "Distributive numerals and distance distributivity in Tlingit (and beyond)". Accessed March 20, 2021. [lingbuzz/001814](http://lingbuzz/001814).
- Champollion, Lucas. 2016. "Covert distributivity in algebraic event semantics". *Semantics and Pragmatics* 9(15): 1–66.
- Choe, Jae-Woong. 1987. *Anti-Quantifiers and a Theory of Distributivity*. PhD Diss., University of Massachusetts Amherst.
- Farkas, Donka. 2015. "Dependent indefinites revisited". Paper presented at the *Workshop on (Co-)Distributivity*, 25-27 February 2015, Paris.
- Henderson, Robert. 2014. "Dependent indefinites and their post-suppositions". *Semantics & Pragmatics* 7 (6): 1-58.
- Henderson, Robert. 2016. "Variation in dependent indefinites". Paper presented at the *Workshop on (Co-)Distributivity*, 11-12 February 2016, Paris.
- Hwang, Sangjin. 2012. *Semantics of measure phrases, measure phrase split, and -ssik in Korean*. Qualifying paper. Accessed March 20, 2021. [http://www.nyu.edu/projects/szabolcsi/hwang\\_ssik\\_and\\_split.pdf](http://www.nyu.edu/projects/szabolcsi/hwang_ssik_and_split.pdf).
- Knežević, Nataša, and Hamida Demirdache. 2018. "Universal Quantification and Distributive Marking in Serbian". In *Linguistic and Cognitive Aspects of Quantification (Studies in Theoretical Psycholinguistics)*, edited by Katalin É. Kiss and Tamás Zétényi, 115-140. Cham: Springer.
- Link, Godehard. 1983. "The logical analysis of plurals and mass terms: A lattice-theoretical approach". In *Meaning, use and interpretation of language*, edited by Reiner Bäuerle, Christoph Schwarze, and Arnim

- von Stechow, 303–323. Reprinted in Link 1998, Chapter 1. Berlin: de Gruyter.
- Link, Godehard. 1987. “Generalized quantifiers and plurals”. In *Generalized quantifiers*, edited by Peter Gärdenfors, 151–180. Reprinted in Link 1998, Chapter 4. Berlin, Germany: Springer.
- Oh, Sei-Rang. 2001. “Distributivity in an event semantics”. In *Proceedings of SALT XI*, edited by Rachel Hastings, Brendan Jackson and Zsófia Zvolensky, 326–345. Ithaca, NY: CLC Publications.
- Oh, Sei-Rang. 2005. *Plurality Markers Across Languages*. PhD diss., University of Connecticut.
- Panaiteescu, Mara. 2018. “Dependent numerals and dependent existentials in Romanian”. *Romance Languages and Linguistic Theory 14. Selected papers from the 46th Linguistic Symposium on Romance Languages (LSRL)*, edited by Lori Repetti and Francisco Ordóñez, 215–229. Amsterdam: John Benjamins.
- Panaiteescu, Mara. 2019. “Distributivity and Additivity: The Case of the Romanian Distributive Marker *câte*”. In *Proceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory*, edited by Anja Gattnar, Robert Hörnig, Melanie Störzer, and Sam Featherston, 305–325. Tübingen: University of Tübingen.
- Rothstein, Susan. 1995. “Adverbial quantification over events”. *Natural Language Semantics* 3 (1): 1–31.
- Schwarzschild, Roger. 1996. *Pluralities*. Vol. 61 (Studies in Linguistics and Philosophy). Dordrecht, Netherlands: Kluwer.
- Szabolcsi, Anna. 2004. “Positive polarity–negative polarity”. *Natural Language and Linguistic Theory* 22 (2): 409–452.
- Winter, Yoad. 1997. “Choice functions and the scopal semantics of indefinites”. *Linguistics and Philosophy* 20 (4): 399–467.
- Winter, Yoad. 2001. *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Cambridge, MA: MIT Press.
- Zimmermann, Malte. 2002. *Boys Buying Two Sausages Each: On the Syntax and Semantics of Distance-Distributivity*. PhD diss., University of Amsterdam.

## Notes

<sup>1</sup> See Winter (1997, 2001) and Champollion (2016) for an overview.

<sup>2</sup> For similar adverbials in English and Tlingit among others see Beck & von Stechow (2007), Braşoveanu & Henderson (2009) for the former and Cable (2013) for the latter.

<sup>3</sup> The notation may be misleading:  $\text{Cov}[x]$  reads “Cov is a cover of  $x$ ”, i.e. Cov is a set containing parts of  $x$ ;  $\text{Cov}(x)$  reads “ $x$  is an element of Cov”.

<sup>4</sup> The same applies to *câte* + MeasP as direct object of a measure verb. See Panaitescu (2019) for examples and analysis.

<sup>5</sup> The term „strength” is used here with respect to the licensor, not the item itself. This is not to be confused with the classification of polarity items themselves in terms of strength in Szabolcsi (2004), in which the stronger the polarity sensitive items, the more licensors it can have (including what I called „weak” ones).

<sup>6</sup> See Balusu (2006) for Telugu cardinal reduplication, Henderson (2014) for Kaqchikel cardinal reduplication, Cable (2013) for Tlingit, Zimmermann (2002) for German *jeweils* etc.

# CHAPTER TEN

## FORTITION IN THE HISTORICAL PHONOLOGY OF MALTESE: TWO CASE STUDIES

ANDREI A. AVRAM

The present paper looks into two instances of fortition in the historical phonology of Maltese: the replacement of interdental fricatives by stops and word-final obstruent devoicing. The records of earlier Maltese examined cover a period ranging from the 15<sup>th</sup> century to the end of the 18<sup>th</sup> century. The data presented and the analysis of the orthography used in the sources, corroborated by metalinguistic comments of contemporary authors, demonstrate that both the replacement of interdental fricatives by stops and word-final obstruent devoicing are rather late developments in the historical phonology of Maltese. It is also shown that both phonological changes are illustrative of lexical diffusion. However, the replacement of interdental fricatives by stops is a case of restructuring, whereas word-final obstruent devoicing is an instance of rule addition and a persistent rule of Maltese.

Keywords: interdental fricatives, word-final obstruents, lexical diffusion

### 1. Introduction

The aims of the present paper are to look at the diachrony of two phonological changes in Maltese: the replacement of interdental fricatives by stops; word-final obstruent devoicing. The changes at issue are both instances of fortition, understood as “a strengthening in the overall force of a sound” (Crystal 2008: 197) or “the increase in magnitude or duration of a gesture” (Bybee 2016: 43). As put by Crystal (2008: 197), “typically, fortition involves the change from a FRICATIVE to a stop [...] or a VOICED to a voiceless sound”.

The corpus of Maltese covers a period ranging from the 15<sup>th</sup> century to the end of the 18<sup>th</sup> century. It consists of texts, wordlists, vocabularies, dictionaries, and lists of place-names and personal names. The texts are Caxaro's *Cantilena* (Wettinger and Fsadni 1968), Buonamico's Sonnet (Cachia 2000), the sermons of Ignazio Saverio Mifsud (Gabra tal-Malti Qadim n.d.), de Soldanis' (1750) grammar, the Christian Catechism (Wzzino 1752), de Soldanis' dialogues (Id-Djalogi ta' de Soldanis n.d.) – written after 1760, and Cannolo's translation of *La Via Sagra* (Cannolo 1796). The lexicographical works consist of Megiser's wordlist (Megiser 1610), Thezan's *Regole per la lingua Maltese* (Cassola 1992), Skippon's wordlist (Skippon 1732), Maius' wordlist (1718), de Sentmenat's Catalan-Maltese vocabulary (Queraltó Bartrés 2003), de Soldanis' dictionary (de Soldanis after 1760), and *Il Mezzo Vocabolario Maltese-Italiano del'700* (Cassola 1996). The place names are from Abela (1647) and Wettinger (1983), while personal names are from Wettinger (1968) and Fiorini (1987–1988).

The timeline of the regressive assimilation of voicing in Maltese is inferred from the orthography used in the sources. Reference is also made to metalinguistic comments by late 18<sup>th</sup>-century authors.

Throughout the text, *t* represents [θ], *d* stands for [ð], *đ* corresponds to [d<sup>h</sup>], and *z* renders [d<sup>l</sup>]. All examples appear in the orthography or transcription system used in the sources. The entries include the original gloss (in Catalan, German, Italian or Latin) and comparative Arabic or Modern Maltese forms. The relevant portions are highlighted in boldface. The following abbreviations are used in the examples: A = Arabic; I = Italian; Mod M = Modern Maltese; S = Sicilian.

The paper is organized as follows. In section 2 it is shown that interdental fricatives are still found in pre-1800 Maltese. Section 3 illustrates the occurrence of word-final obstruent devoicing in earlier records of Maltese. Section 4 discusses the findings and their implications.

## 2. Interdental fricatives

Caxaro's *Cantilena* (Wettinger and Fsadni 1968, Cohen and Vanhove 1991) is the earliest extant text in Maltese, dated c. 1450. It contains the following occurrence of \**t*, rendered with the digraph <th>:

- (1) *nichadithicum* 'I shall tell you', cf. A *ħadda* *ħ* 'to relate, to report'

Interdental fricatives are attested in 16<sup>th</sup>-century sources. In a place-name recorded in a 1501 notarial document the digraph <th> corresponds to etymological \**t̪*:

- (2) *gebel labiath*<sup>1</sup> ‘the white rock, the white hill-side’, cf. A *ʔabyaḏ*

The occurrence of interdental fricatives is better represented in Megiser’s wordlists, collected in Malta, in 1588 (Cowan 1964, Cassola 1987–1988, Avram 2014). In the first, a 10-word list with glosses in Latin, published in 1603 (Cassola 1987–1988), the letter <f> corresponds to etymological \**t̪*:

- (3) a. *fne* ‘duo’, cf. A *iṣṣnāni*  
 b. *fliesan* ‘tres’, cf. A *ṭalāṭa*

Etymological \**t̪* is also rendered with <s>:

- (4) *fliesan* ‘tres’, cf. A *ṭalāṣa*

The second, a 121-word list (Megiser 1610: 9-14) with glosses in German, contains several forms attesting to the occurrence of *t̪*, which is transcribed in different ways. For instance, in the examples below etymological \**t̪* is transcribed with <f>:

- (5) a. *Fne* ‘Zwey’, cf. A *iṣṣnāni*  
 b. *Fliesan* ‘Drey’, cf. A *ṣalāṭa*

In the following form <s> corresponds to etymological \**t̪*:

- (6) *fliesan* ‘Drey’, cf. A *ṭalāṣa*

The digraph <sf> also corresponds to etymological \**t̪*:

- (7) a. *Sfniema* ‘Achte’, cf. A *ṣamāniya*  
 b. *Sfremi* ‘Achzig’, cf. A *ṣamānīn*

Finally, in the form in (8) it is <h> which corresponds to etymological \**t̪*:

- (8) *Flehmi*a ‘Dreyhundert’, cf. A *ṭalāṣ miya*

Since <f>, <s>, <sf> and <h> all render voiceless fricatives, these spellings are indicative of the occurrence of the interdental voiceless fricative *t̪*, for which German spelling has no equivalent. In addition, one word suggests that its voiced counterpart *d̪* still occurred as well:

(9) *Veheb* ‘Gold’, cf. A *Ṣahab*

The use of <v>, which renders a voiced fricative, is most probably an attempt at transcribing the interdental voiced fricative *d̪*.

A first piece of evidence for the occurrence of interdental fricatives in 17<sup>th</sup>-century Maltese is provided by the place-names recorded by Abela (1647). In the transcription system employed, the digraph <th> consistently corresponds to etymological *\*t̪*:

- (10) a. *Ghar Buthomna* ‘The Cave of the Measure of Corn’<sup>2</sup>, cf. A *Ṣumna* ‘measure of corn’  
 b. *Ben Varrath* ‘figlio dell’ Herede’, cf. A *warīt* ‘heir’  
 c. *Ghar el MethKub* ‘grotta portugiata, ò forata’, cf. A *maṢqūb* ‘perforated’

The digraph <dh> corresponds to etymological *d̪*:

- (11) a. *Bir el dheeb* ‘pozzo d’oro’, cf. A *dahab*  
 b. *Dhoccara* ‘wild fig’<sup>3</sup>, cf. A *Ṣakkara* ‘wild fig’

Moreover, the same digraph also corresponds to etymological *d̪*. Consider the following examples:

- (12) a. *Blata el baydha* ‘rocca bianca’, cf. A *baiḏā?*  
 b. *Hal DheEIF* ‘Casale del macilente, o debole’, cf. A *ḏaṣīf*

Consider next Thezan’s dictionary (Cassola 1992), written by 1647. The author uses a mixed orthography, consisting of Latin and Arabic characters. The fact that the latter include <ḏ> / <ḥ> and <ḏ̣>, which represent the interdental fricatives *t̪* and *d̪* respectively, therefore constitutes convincing evidence of their survival in 17<sup>th</sup>-century Maltese. This is further reinforced by the large number of entries which exhibit interdental fricatives. Reflexes of etymological *\*t̪*, transcribed with <ḏ> or <ḥ>, are amply documented in all positions, word-initially (13a-e), word-medially (13f-h), and word-finally (13i):



- (13) a.  $\dot{\text{z}}$  *elieta* ‘tre’, cf. A *Ḥalāta*  
 b.  $\dot{\text{z}}$  *emienia* ‘otto’, cf. A *ṭamāniya*  
 c.  $\dot{\text{z}}$  *eum* ‘aglio’, cf. A *Ḥūm*  
 d.  $\dot{\text{z}}$  *omona* ‘tomena, misura’, cf. A *Ḥumna* ‘a measure of corn’  
 e.  $\dot{\text{z}}$  *eni* ‘duplicare’, cf. A *Ḥanā*  
 f. *ak*  $\dot{\text{z}}$  *ar* ‘piu, assai piu’, cf. A *akḤar*  
 g. *me*  $\dot{\text{z}}$  *kub* ‘pertuzato. forrato’, cf. A *maḤqūb*  
 h. *tḥade*  $\dot{\text{z}}$  *r*<sup>4</sup> ‘ragionare, discorrere, favellare, parlare’, cf. A *taḥaddaḤt*  
 i. *mo*  $\dot{\text{z}}$  *ere* *trāḥmi* A cf. aratro’,  $\dot{\text{z}}$ <sup>5</sup>

Similarly, reflexes of etymological \**d*, transcribed with < $\dot{\text{z}}$ > are attested in word-initial (14a-b), word-medial (14c) and word-final (14d) position:

- (14) a.  $\dot{\text{z}}$  *eḥeb* ‘oro’, cf. A *ḍahab*  
 b.  $\dot{\text{z}}$  *eil* ‘falda di camiscia ò daltro’, cf. A *Ḥail* ‘garment’  
 c. *ghe*  $\dot{\text{z}}$  *ieb* ‘bugiardo’, cf. A *kaddāb*  
 d. *emb*  $\dot{\text{z}}$  *ī* ‘vino’, cf. A *nabīḍ*

As shown by the examples under (15), the letter < $\dot{\text{z}}$ > also corresponds to etymological \**d*, word-initially (15a-c), word-medially (15d-e) and word-finally (15f):

- (15) a.  $\dot{\text{z}}$  *arab* ‘ferire’, cf. A *ḍarab*  
 b.  $\dot{\text{z}}$  *au* ‘luce, lume’, cf. A *ḍau?*  
 c.  $\dot{\text{z}}$  *om* ‘cogliere’, cf. A *ḍamma*  
 d. *fe*  $\dot{\text{z}}$  *a* ‘argento non monetato’, cf. A *fiḍḍa*  
 e. *me*  $\dot{\text{z}}$  *erub* ‘ferito’, cf. A *maḍrūb*  
 f. *mar*  $\dot{\text{z}}$  *ī* ‘malatia’, cf. A *maḍī*

Finally, in a number of entries < $\dot{\text{z}}$ > also corresponds to etymological \**Ḥ*. Consider the forms below:

- (16) a.  $\dot{\text{z}}$  *an* ‘credere parere pensare’, cf. *Ḥanna*  
 b.  $\dot{\text{z}}$  *el* ‘ombra’, cf. A *Ḥill*  
 c.  $\dot{\text{z}}$  *elam* ‘oscuro’, cf. A *Ḥalām* ‘darkness’  
 d. *a*  $\dot{\text{z}}$  *am*<sup>6</sup> ‘osso’, cf. A *Ḥam*  
 e. *ne*  $\dot{\text{z}}$  *af* ‘nettare’, cf. A *naḤḤaf*

The occurrence of interdental fricatives is documented in 18<sup>th</sup>-century sources as well. Consider first de Sentmenat's Catalan-Maltese vocabulary (Queraltó Bartrés (2003), dating from around 1750. In the entries reproduced below the digraph <th> corresponds to etymological \*t̪:

- (17) a. *Theminin* 'Vuiytanta', cf. A *Ṣamānīn*  
 b. *Theuma* 'all', cf. A *Ṣūm*  
 c. *Thimienia* 'vuyt', cf. A *Ṣamāniya*  
 d. *Thòmena* 'Mesura de Blat', cf. A *Ṣumna*

The digraph <dh> corresponds to etymological \*d̪ in the following forms:

- (18) a. *Dhehhheb* 'or' cf. A *Ṣahab*  
 b. *Dhèil* 'camisa de dona', cf. A *Ṣail* 'garment'

Moreover, in one entry the sequence <dhh> corresponds to etymological \*d̪d̪:

- (19) *Phidhhhá* 'Plat', cf. A *fiḏḏa*

Consider next de Soldanis' (1750) grammar and dictionary of Maltese. The digraph <th> corresponds to the Arabic letter *ṯ*, which suggests that the consonant is the voiceless interdental fricative t̪:

- (20) *Thielet* 'il terzo', cf. A *Ṣālit* 'third'

Moreover, de Soldanis (1750: 71) underscores, albeit inadequately, the fact that <th> stands for a sound which differs from [t], writing "The *TH th* Alquanto duro, ma aspirato". The voiced interdental fricative is also attested. In de Soldanis' (1750) alphabet, the specially designed letters <Ḑ>, <ḑ> correspond to etymological \*d̪, as in the following examples:

- (21) a. *Ṣeep* 'oro', cf. A. *Ṣahab* 'gold'  
 b. *Ṣeil* 'camiscia dimezzata', cf. A *Ṣail* 'garment'

The nature of the consonant transcribed with <ḑ> can also be inferred from the comments by de Soldanis (1750: 72-73) on the letters <Ḑ>, <ḑ>: "ḑ Dal *D d* Questo col punto sopra diviene blesa". The term *blesa* 'with a lisp' shows that the consonant at issue is phonetically realized as a interdental fricative.

Interdental fricatives are last attested in two other works by de Soldanis. In his four-volume dictionary (de Soldanis after 1760), the

specially designed letter <î> corresponds to etymological \**t*, as in the following examples:

- (22) a. *mkaṡṡar* ‘multiplied’, cf. A *mukaṡṡar*  
 b. *mṡennia* ‘repeated’, cf. A *maṡniya* ‘doubled’  
 c. *ṡielet* ‘third’, cf. A *ṡaliṡ* ‘third’

The letter <ḍ> corresponds both to etymological \**d* and to etymological \**ḍ*. This is illustrated in (23) and (24) respectively:

- (23) a. *ṡeeb* ‘gold’, cf. A. *ṡahab*  
 b. *ṡeen* ‘mind’, cf. A *ṡihn*
- (24) a. *ṡamma* ‘collection’, cf. A *ḍamma*  
 b. *ṡaul* ‘light’, cf. A *ḍau?*  
 c. *ṡarba* ‘time’, cf. A *ḍarba*

Further examples of voiced interdental fricatives are found in de Soldanis’ dialogues (Id-Djalogi ta’ de Soldanis n.d.). Here again <ḍ> corresponds either to etymological \**d*, as in (25), or to etymological \**z*, as in the forms under (26):

- (25) *ṡaul* ‘light’, cf. A *ḍau?*
- (26) a. *ṡellu* ‘his shadow’, cf. A *ṡill* ‘shadow’  
 b. *ṡlam* ‘darkness’, cf. A *ṡalām*

### 3. Word-final obstruent devoicing

Word-final obstruent devoicing is not attested in Caxaro’s *Cantilena*, the earliest known Maltese text. However, personal names recorded in 15<sup>th</sup>-century notarial documents include extremely rare cases of obstruents devoiced word-finally (see Avram 2012):

- (27) a. *Muhamet*, cf. A *Muḥammad*  
 b. *aius*, cf. A *ṡağūz*

Note that in (27b) <s> may stand for [z], since <z> would represent [ts] or [dʒ] according to the orthographic conventions of Sicilian. If so, (27b) does not illustrate word-final obstruent devoicing.

Convincing instances of word-final obstruent devoicing are first found in Megiser's word-list, collected in 1588 (see Megiser 1606). These consist of voiceless reflexes of etymological voiced stops, as in (28a-d), and respectively fricatives, as in (28e-f):

- (28) a. *Ecnep* 'Trauben', cf. A *ʃinab*  
 b. *Tajep* 'Schöne heutere Zeit', cf. A *ʃayyib* 'good'  
 c. *it* 'Hand', cf. A *yad*  
 d. *Guart* 'Rosen', cf. A *ward*  
 e. *Embit* 'Wein', cf. A *nabīʃ*  
 f. *Chops* 'Brodt', cf. A *ħubz*

The evidence provided by Megiser (1606) is reliable since, even though he was a native speaker of German, a language which also exhibits word-final obstruent devoicing, he used both <p> and <b> for the transcription of word-final bilabial stops. It can therefore be assumed that Megiser's transcriptions capture variation in the phonetic realization of word-final obstruents.

17<sup>th</sup>-century records of Maltese provide much more evidence for the occurrence of word-final obstruent devoicing. The majority of the relevant entries in Thezan's dictionary (Cassola 1992), written by 1647, still have voiced obstruents word-finally. However, all types of obstruents, i.e. stops (29a-b), fricatives (29c) or affricates (29d) may undergo devoicing:

- (29) a. *ġarp*<sup>7</sup> 'ponente', cf. A *garb*  
 b. *ermiet* 'ashes', cf. A *ramād*  
 c. *ʃabes*<sup>8</sup> 'saltare', cf. Mod M *qabeʒ* 'to jump'  
 d. *aɣa*<sup>9</sup> 'avorio dente d'elefante', cf. A *ʃāʃ*

Devoiced word-final obstruents are also found in a few lexical items of Romance origin:

- (30) a. *boros* 'borsa' [plural of *borza*], cf. S *burza*<sup>10</sup>  
 b. *preies* 'preda, presa' [plural of *preza*]', cf. I *presa*

The place-names recorded by Abela (1647) include forms with both voiced and devoiced obstruents in word-final position. As shown below, the latter exhibit voiceless reflexes of voiced etymological stops (31a-b) and fricatives (31c):

- (31) a. *Mitahlep* 'luogo, oue si mungeua il latte', cf. A *√ħlb*

- b. *Kibur elihut* ‘Sepulchri de’ Giudei’, cf. A *al-yahūd*
- c. *Redùm el Bies* ‘dirocato del falcone’, cf. A *bāz* ‘falcon’

Only a minority of the relevant entries in Skippon’s wordlist, collected in 1664 and published later (Skippon 1732), illustrate the occurrence of word-final obstruent devoicing. The forms include several Arabic-derived words, as in (32), and the Romance loanword in (33):

- (32) a. *tachsep* ‘cogitare’, cf. A *taḥsab* ‘you think’
  - b. *raat* ‘tonitru’, cf. A *raḥad* ‘to thunder’
  - c. *akbes* ‘saltare’, cf. Mod M *qabež* ‘to jump’
- (33) *perikulus* ‘pericolosum’, cf. S *periculusu*<sup>11</sup>

It should be mentioned that, since Skippon consistently uses the letter <z> to represent [z], the <s> in (32c) and (33) stands for [s].

In the last 17<sup>th</sup>-century source examined, Buonamico’s *Sonnet* (Cachia 2000: 18), dated 1672, two out of the five relevant forms have a devoiced obstruent in word-final position:

- (34) a. *art* ‘earth’, cf. A *ʔard*
- b. *bart* ‘cold’, cf. A *bard*

Word-final obstruent devoicing is amply documented in a variety of 18<sup>th</sup>-century texts, by both foreigners or native speakers of Maltese. In Maius’ (1718) short list of Maltese words, three forms exhibit devoiced obstruents in word-final position:

- (35) a. *it* ‘hand’, cf. A *yad*
- b. *Guart* ‘rose’, cf. A *ward*
- c. *Hops* ‘bread’, cf. A *ħubz*

The occurrence of these forms cannot be attributed to the fact that Maius was a native speaker of German, in which word-final obstruent devoicing is also found. Indeed, Maius also employs <b> word-finally, which is suggestive of variation in the phonetic realization of obstruents in this position.

The majority of the relevant forms in the sermons by Mifsud (Ġabra tal-Malti Qadim n.d.), dating from the period 1739–1746, display devoiced word-final obstruents. Consider the following voiceless reflexes

of voiced stops (36a-b), fricatives (36c) and affricates (36d) in Arabic-derived words:

- (36) a. *taiep* ‘good’, cf. A *ṭayyib*  
 b. *iocot* ‘[he] sits’, cf. Mod M *joqgħod*  
 c. *hops* ‘bread’, cf. A *ḥubz*  
 d. *uic* ‘face’, cf. A *waḥh*

There is also a Romance loanword in which a stop is devoiced word-finally:

- (37) *tart* ‘late’, cf. I *tardi*<sup>12</sup>

Devoiced and voiced word-final obstruents are approximately evenly distributed among the entries in marquis de Sentmenat’s Catalan-Maltese vocabulary (Queraltó Bartrés 2003), dated c. 1750. Consider the following Arabic-derived words with voiceless reflexes of voiced stops (38a-b), fricatives (38c-d) and affricates (38e):

- (38) a. *Hhalip* ‘Llet’, cf. A *ḥalīb*  
 b. *Bart* ‘Fret’, cf. A *bard*  
 c. *imbit* ‘Vi’, cf. A *nabīḥ*  
 d. *Hhops* ‘Pa’, cf. A *ḥubz*  
 e. *Mahhmux*<sup>13</sup> ‘Brut’, cf. A *maḥmūḥ* ‘rotten’

Word-final obstruent devoicing is also attested in a Romance loanword:

- (39) *Txerph*<sup>14</sup> ‘Ciervo’, cf. S *cervu*, I *cervo*<sup>15</sup>

Since voiced obstruents are also found word-finally, this points to variation, rather than the influence of the author’s native language, Catalan, which also has final devoicing.

Another 18<sup>th</sup>-century source documenting the occurrence of word-final obstruent devoicing is the grammar and short dictionary of Maltese by de Soldanis (1750). This source is particularly important. Firstly, forms with devoiced word-final obstruents outnumber those with voiced ones. Consider the Arabic-derived words below:

- (40) a. *Qtiep* ‘libro, volume’, cf. A *kitāb*  
 b. *Takap* ‘forò, o traspasò’ cf. A *ṭakab*  
 c. *l’art* ‘la terra’, cf. A *al-ʔarḍ*

- d. *uart* ‘rosa’, cf. A *ward*
- f. *Saiet* ‘pescatore’, cf. A *ṣayyād*
- g. *Hhaps* ‘ritegno’, cf. A *ḥabs*

Forms with devoiced word-final obstruents also include Romance loanwords:

- (41) a. *squt* ‘scudo’, cf. S *scudu*, I *scudo*<sup>16</sup>
- b. *supperf* ‘rigido’, cf. I *superbo*<sup>17</sup>

Secondly, de Soldanis (1750) provides uncontroversial evidence of variation in the phonetic realization of obstruents in word-final position. This is pointed out on several occasions, both in the grammar (42a) and in the dictionary (42b-c):

- (42) a. *Bieb* o *Biep* ‘porta’, cf. A *bāb*
- b. *tajeb*, o *tajep* ‘buono’, cf. A *ṭayyib*
- c. *Rmiēt*, *Ramed* ‘ashes’, cf. Ar. *ramād*

Furthermore, several other forms, such as the ones in (43), illustrate variation:

- (43) a. *Mqareb* ‘rigid man’ vs. *mḡārep* ‘rigid’, cf. Mod. M. *mḡareb*
- b. *ghand* vs. *ghant* ‘of’, cf. Ar. *ʕind*

In Wzzino’s (1752) translation of the Christian Catechism devoiced word-final obstruents are in the minority. However, these are found in reflexes of all types of obstruents – voiced stops (44a-b), fricatives (44c-e), and affricates (44f):

- (44) a. *itlop* ‘[he] asks’, cf. A *yaṭlub*
- b. *uiet* ‘valley’, cf. A *wād-*
- c. *imbit* ‘wine’, cf. A *nabīṣ*
- d. *ghāzis* ‘beloved’, cf. A *ʕazīz*
- e. *uisch*<sup>18</sup> ‘face’, cf. A *waḡh*

De Soldanis’ dialogues (Id-Djalogi ta’ de Soldanis n.d.) is yet another 18<sup>th</sup>-century source in which forms with devoiced word-final obstruents outnumber those in which voiced ones are found. The former include the examples under (45):

- (45) a. *i gip* ‘[he] brings’, cf. Mod M *iḡib*

- b. *trap* ‘dust’, cf. A *turāb*
- c. *art* ‘earth’, cf. A *ʔard*
- d. *uuihhet* ‘one’, cf. A *wāḥid*
- e. *uec* ‘face’, cf. A *waḥh*

Significantly, there is also evidence of both inter- and intra-speaker variation in the phonetic realization of word-final obstruents, as shown in (46) and (47), respectively; this occurs either in the same word or in different forms in the paradigm of a word:

- (46) a. *t okghod* vs. *t okghot* ‘[you] sit’, cf. Mod. M. *toqgħod*
- b. *t ahhseb* ‘[you] think’ vs. *n-ahhsep* ‘[I] think’, cf. Mod. M. *taħseb, naħseb*
- (47) a. *e sib* ‘[he] finds’ vs. *sap* ‘[he] found’, cf. Mod. M. *isib, sab*
- b. *embaghad* vs. *embaghat* ‘then’, cf. Mod M *mbagħad*

A later source, *Il Mezzo Vocabolario Maltese-Italian del '700* (Cassola 1996), written by 1775, rather surprisingly contains very few forms in which obstruents are devoiced word-finally. Note, however, that these include reflexes of all types of voiced obstruents, stops (48a), fricatives (48b), and affricates (48c):

- (48) a. *Miliet* ‘Natale di nostro Signore Gesù Cristo’, cf. A *milād*
- b. *nekkês* ‘pungere’, cf. A *nakaz*
- c. *uete*<sup>19</sup> ‘faccia’, cf. A *waḥh*

The last 18<sup>th</sup>-century source examined is Cannolo’s (1796) translation of *La Via Sagra*. While the total number of relevant items is small, it is significant that the ratio between forms with devoiced word-final obstruents and those with voiced ones is approximately 2 to 1. The former include both Arabic-derived words (49), and a Romance loanword (50):

- (49) a. *l’Hut* ‘the Jews’, cf. Mod M *Lhud*
- b. *hhuejecc* ‘things’, cf. Mod M *ħwejjegħ*
- (50) *schont* ‘according to’, cf. I *secondo*<sup>20</sup>



#### 4. Discussion and conclusion

Maltese certainly had interdental fricatives in its earlier stages, contra e.g. Agius (1996: 272), who writes that “Maltese does not have the interdental fricative /d̪/ and historically may have never experienced this articulation”. The relative chronology of the replacement of interdental fricatives by stops also needs to be revised. According to Cohen (1967, 1967), the chronology is as follows: stage (i)  $d$  ( $d̪$ )  $d̪ z > d d$ ; stage (ii)  $d̪ d̪ > d$ . However, the data presented in section 2 show that  $d̪$  is also the reflex of etymological  $*d̪$  and  $*z̪$ . The relative chronology is therefore: stage (i)  $d̪ d̪ z̪ > d d̪ z̪$ ; stage (ii)  $d̪ d̪ z̪ > d d̪$ ; stage (iii)  $d̪ d̪ > d$  (see also Avram 2014, 2016). Finally, the replacement of interdental fricatives by stops appears to be an internal development. As for word-final obstruent devoicing, this may be the result of imperfect L2 acquisition of the phonology of Maltese by Sicilian- and Italian-speaking immigrants to Malta, hence, a contact-induced change (Avram 2017a).

Consider next the two phonological changes in the light of a possible connection between Maltese and Sicilian Arabic<sup>21</sup>. The occurrence of interdental fricatives in early Maltese is compatible with the hypothesis according to which (a variety of) Sicilian Arabic developed into the former. However, it is incompatible with Agius’ (1996, 2007) claim that Maltese originated in a creolized form of Arabic presumably spoken in Arab-occupied Sicily. As is well known, fricative interdentals are cross-linguistically marked consonants and are typically not preserved in pidgins and creoles (Avram 2014). As far as word-final obstruent devoicing is concerned, the case for (a variety of) Sicilian Arabic as the possible ancestor of Maltese is considerably weakened by the fact that the process also occurs, even if inconsistently, in the former. This does not therefore accord with the late emergence of word-final obstruent devoicing in the Maltese (Avram 2017b).

Both phonological changes at issue are late developments, having essentially run their full course towards the end of the 18<sup>th</sup> century (Avram 2012, 2014, 2016, 2017a, 2020). Note that the replacement of interdental fricatives occurs much later than claimed in the literature. Cohen (1966: 13), for example, states that “la confusion des deux articulations [dentales et interdentales] peut [...] remonter à un stade pré-maltais”. Similarly, Cohen (1967: 168) writes that “on peut attribuer à un stade pré-maltais la confusion des dentales et des interdentales”. According to Vanhove (1994: 170), “l’absence des interdentales remonte à la période pré-historique du maltais”. More recently, Vanhove (2000: 189) concludes that “l’absence

des interdentes [...] semble bien être un héritage direct de la variété d'arabe maghrébin I provient le maltais”.

The data presented in sections 2 and 3 suggest that both the replacement of interdental fricatives by stops and word-final obstruent devoicing are cases of lexical diffusion. The central claim of lexical diffusion theory (see Chen and Wang 1975, McMahon 1994: 50, Phillips 2015, Bybee 2016: 39, Burridge and Bergs 2017: 163-167, among others) is that sound change is phonetically abrupt, but lexically gradual, i.e. not all words are affected at the same time. In the initial stage, then, sound change does not affect all relevant lexical items, but occurs only in a small subset of the potential targets. For a certain period of time there will be variation, characterized by the coexistence of two competing phonetic realizations – the initial one and the new one resulting from the sound change. At a still later stage, the sound change spreads both to other lexical items as well as to other speakers. Finally, the sound change may extend to all relevant words.

Consider first the replacement of interdental fricatives by stops. As seen in Table 1 below, interdental fricatives persist for a longer period in some words:

Caxaro (c. 1450)	Megiser (1610)	Abela (1647)	Thezan (1647)	de Sentmenat (c. 1750)	de Soldanis (1750)	de Soldanis (after 1760)
<b>Reflexes of etymological *ʒ</b>						
<i>nichadithicum</i>	<i>Fliesan</i> <i>Sfniema</i>	<i>Buthomna</i> <i>MethKub</i>	تadeث ‡elieta ‡ emienia ‡eum ‡omona me ‡kub	<i>Thimienia</i> <i>Theuma</i> <i>Thòmena</i>	<i>thielet</i>	‡ielet
<b>Reflexes of etymological *ʒ</b>						
	<i>veheb</i>	<i>dheeb</i> <i>Dhoccara</i>	‡e‡eb ‡eil ‡akkar	<i>dhehhheb</i> <i>Dhèil</i>	‡eep ‡eil	‡eeb

Reflexes of etymological *ǧ				
	ǧ arab	ǧarba		
	ǧ au		ǧaul	ǧaul
	ǧ om		ǧamma	
baidha	a ǧbia ǧ			
	feǧ ǧ a	Phidhhhá		
Reflexes of etymological *ǧ				
	ǧ el			ǧell-
	ǧ elam			ǧallam
				ǧlam

**Table 1. Persistence of interdental fricatives**

There is also evidence of variation, as predicted by lexical diffusion theory. On the one hand, no interdental fricatives are attested in e.g. Skippon’s wordlist collected in 1664 (Skippon 1732), Buonamico (1672), Mifsud (1739–1746), Wzzino (1752), and *Il Mezzo Vocabolario Maltese-Italiano del ’700* (by 1775). On the other hand, interdental fricatives are possibly preserved until later in the Gozitan dialect, which is demonstrably more archaic (Vella 2013). Variation in the occurrence of interdental fricatives is further confirmed by metalinguistic comments by contemporary authors. For instance, in a letter to de Soldanis, dated September 27, 1749, G. Cachia states about *t̪* that “nella pronuncia della lettera *t* e *th*, non vi è gran differenza”<sup>22</sup>. This shows that by that time, there already was confusion of the interdental fricatives and the corresponding stops. The clearest evidence of the gradual spread of fortition of the interdental fricatives, both through the lexicon and in the speech community, is found in Vassalli (1796). In the “Discorso preliminare” to his dictionary, Vassalli (1796: XXIX, f.n. 11) writes that, unlike Arabic, which has *t*, *t̪*, *t̪*, *d*, *d̪*, and *z*, “noi [= the Maltese] non distinguiamo, fuori del *The* e del *Dhal*”, i.e. *t̪* and *d̪* as opposed to *t* and *d*. While *t̪* and *d̪* still existed, Vassalli (1796: XXIX, f.n. 11) adds that these are “molto raramente e tra pochi usati”, i.e. the occurrence of interdental fricatives was confined to few speakers in that period. Evidence for the gradual spread of fortition of the interdental fricatives through the lexicon of Maltese is also provided; consider Vassalli’s (1796: 62) comments *De littera T*: “*Hossha hu bħal fyl-Latīn u fyt-Taljān, ǧhād xi drābi tkūn mtemtma: kīf nystgħu narav fl’aħħar ta dān yl kt̪yb, fejn jynsābu yl kelm̪y̪t kollha mtemtm̪īn*”<sup>23</sup>.

Word-final obstruent devoicing also appears to have been implemented slowly and gradually, with some words being first affected earlier than others, as illustrated in Table 2 below:

Megiser (1588)	Thezan (by 1647)	Skippon (1664)	Buonamico (1672)	Maius (1718)	Mifsud (1739–1746)	de Sentmenat (c. 1750)	de Soldanis (1750)	Wzzino (1752)	de Soldanis (after 1760)
	<i>bard</i>	<i>bard</i>	<i>bart</i>			<i>bart</i>			
<i>ecnep</i>	<i>eċneb</i>					<i>hhhenep</i>	<i>ghenep</i>		
<i>veheb</i>	<i>veċeb</i>	<i>deheb</i>		<i>Deeb</i>	<i>deep</i>	<i>dhehhheb</i>	<i>ċeep</i>		<i>ċeeb</i>
<i>embit</i>	<i>embiċ</i>					<i>imbit</i>		<i>imbit</i>	<i>embit</i>
		<i>ug</i>	<i>uecce</i>		<i>uic</i>	<i>hhhutx</i>		<i>uisch</i>	<i>uec</i>

**Table 2. Spread of word-final obstruent devoicing (from Avram 2020)**

In confirmation of the scenario posited by lexical diffusion theory, there is considerable variation in the occurrence of word-final obstruent devoicing. As shown in Table 3, forms with voiced obstruents in word-final position are still found in records from the second half of the 18<sup>th</sup> century, even though the same lexical items exhibit word-final obstruent devoicing in earlier sources:

Megiser (1588)	Thezan (by 1647)	Skippon (1664)	Maius (1718)	Mifsud (1739–1746)	de Sentmenat (c. 1750)	Wzzino (1752)	de Soldanis (1750)	de Soldanis (after 1760)
	<i>bieb</i>		<i>bieb</i>	<i>biep</i>	<i>bieb</i>		<i>biep</i>	<i>bieb</i>
							<i>bieb</i>	
	<i>salib</i>			<i>salip</i>	<i>salip</i>	<i>salip</i>	<i>salip</i>	<i>salib</i>
				<i>gdit</i>	<i>xdid</i>	<i>gdit</i>	<i>gidid</i>	<i>gdid</i>
	<i>oċod</i>	<i>okod</i>		<i>iocot</i>	<i>okhhhôt</i>	<i>nokghod</i>		<i>okghot</i>
								<i>okghod</i>
<i>chops</i>	<i>ċobs</i>		<i>hops</i>	<i>hops</i>	<i>hhops</i>	<i>ĥhops</i>	<i>chops</i>	<i>chobs</i>
	<i>ċobz</i>						<i>hhops</i>	
	<i>zeutieg</i>			<i>suiec</i>	<i>zeuhuetx</i>	<i>zueig</i>		

**Table 3. Voiced ~ devoiced word-final obstruents (from Avram 2020)**

Also, by the end of the 18<sup>th</sup> century word-final obstruent devoicing is explicitly mentioned for the first time. Several passages in Vassalli's (1791) grammar of Maltese refer to the devoicing of obstruents in word-final position. Suffice it to quote the following, which is the first statement of the rule of word-final obstruent devoicing in Maltese: “litterae B D Γ [...] & Z in fine dictionum sonos proprios amittunt & pronunciatum ac si essent P T Q<sup>24</sup> [...] & S” (Vassalli 1791: 95).

It must be acknowledged that several factors make it impossible to follow in more detail the processes of replacement of the interdental fricatives by stops and of word-final obstruent devoicing. For instance, 15<sup>th</sup>- and 16<sup>th</sup>-century Maltese is very poorly documented. Also, the available records of earlier Maltese are certainly representative of several dialectal varieties. While “we do not know whether Megiser or Skippon took those words from the city or from the countryside” (Cardona 1997: 22), it appears that Thezan's dictionary reflects the variety of 17<sup>th</sup>-century Maltese spoken in the area of Valletta (Hull 1994: 394), and de Soldanis' works represent the dialect of Gozo (Vella 2013). Moreover, the records examined in sections 2 and 3 are illustrative of a variety of genres. Last, but not least, the task is not made any easier by the inconsistencies in the transcriptions used in the sources.

Finally, the two processes of fortition have different effects on the phonology of Maltese. The replacement of interdental fricatives by stops is a case of restructuring (in the sense of e.g. King 1969, Uguzzoni 1977). Consequently, interdental fricatives are no longer part of underlying representations in Modern Maltese. Word-final obstruent devoicing, however, is a case of rule addition (King 1969, Uguzzoni 1977, McMahon 1994, Drescher 2015), i.e. a phonological change which does not trigger the modification of underlying representations (Avram 2020). As shown by e.g. (Borg 1997: 250), in Modern Maltese “voicing contrasts are normally restricted to underlying representations”. The fact that “voiced obstruents do not occur word-finally [is] a result of the word-final obstruent devoicing rule” (Borg and Azzopardi-Alexander 1997: 307)<sup>25</sup>. Word-final obstruent devoicing operates as a synchronic rule of Modern Maltese, applying both to the native stock of lexical items and to loanwords<sup>26</sup> and is thus a persistent rule (in the sense of Cser 2015) of Maltese phonology.

## Sources

Abela, Giovanni Francesco. 1647. *Della descrizione di Malta isola nel mare siciliano. Con le sue antichità, ed altre notizie*. Malta: Paolo Bonacota.

- Cassola, Arnold. 1992. *The Biblioteca Vallicelliana Regole per la lingua Maltese. The Earliest Extant Grammar and Dictionary of the Maltese Language*. Valletta: Said International.
- Cassola, Arnold. 1996. *Il Mezzo Vocabolario Maltese-Italiano del '700*. Valletta: Said International.
- Cannolo, Giuseppe Martino. 1796. *La Via Sagra. Tradotta in Maltese da Giuseppe Martino Cannolo*. Malta:s.n..
- Ġabra tal-Malti Qdim. n.d.  
<http://mlrs.research.um.edu.mt/resources/gabrastorika>.
- Id-Djalogi ta' de Soldanis. n.d.  
[http://malti.skola.edu.mt/kotba/djalogi\\_desoldanis.pdf](http://malti.skola.edu.mt/kotba/djalogi_desoldanis.pdf).
- Maius, Ioannes H. 1718. *Specimen linguae punicae in hodierna Melitensiun superstitis*. Marburg: Muller.
- Megiser, Hieronymus. 1606. *Propugnaculum Europe. Wahrhafte/Eigentliche und Ausführliche Beschreibung/der viel und weitberühmten Afrikanischen Insul Malta*. Leipzig: Henning Grossen des Jüngern.
- Queralto Bartrés, Alexandre. 2003. *Un vocabulari català – maltès manuscript del segle XVIII. Edició, transcripció i estudi*. Barcelona: PPU.
- de Soldanis, Giovanni Pietro Francesco A. 1750. *Nuova scuola di grammatica per agevolmente apprendere la lingua punica maltese, aperta agli studenti maltesi, e forestieri abitanti in Malta*. Rome: Salomoni.
- de Soldanis, Giovanni Pietro Francesco A. after 1760. “Ġamma tal Kliem Kartaginis mscerred fel fom tal Maltin u Ghaucin”. Ms 143/I-IV, National Library of Malta.
- Skippon, Philip. 1732. *An Account of a Journey Made thro' Part of the Low-Countries, Germany, Italy, and France*. London: s.n.
- Vassalli, Michaelis Antonii. 1791. *Mylsen Phoenico-Punicum sive Grammatica Melitensis*. Rome: Antonio Fulgoni
- Vassalli, Michaelis Antonii. 1796. *Ktýb yl Klým Málti. Mfysser byl-Latin u byt-Taljan. Lexicon Melitense-Latin-Italum. Vocabolario maltese. Recato nelle lingue Latina e italiana*. Rome: Antonio Fulgoni.
- Wettinger, Godfrey and Fsadni, Michael. 1968. *Peter Caxaro's Cantilena. A Poem in Medieval Maltese*. Malta: s.n..
- Wzzino, Francesco. 1752. *Taġħlim nistrani. Migiub fil Ġharbi bl'ordini ta sant. mem. ta Paulu V. Min gdid stampat fit-Talian mis Sacra Congregazioni tal Propaganda Fidi, u' migiub issa fil Malti*. Rome: Arcangelo Casaletti.

## References

- Agius, Dionisius A. 1996. *Siculo Arabic*. London and New York: Kegan Paul International.
- Agius, Dionisius A. 2007. “Who Spoke Siculo-Arabic?”. In *XII Incotro italiano di linguistica camito-semitica (Afroasiatica)*. Atti, edited by Marco Moriggi, 25-33. Soveria Mannelli: Rubbettino.
- Aquilina, Joseph. 1987a. *Maltese – English Dictionary*, vol. one, A - L. Valletta: Midsea Books.
- Aquilina, Joseph. 1987b. *Maltese – English Dictionary*, vol. two, M - Z. Valletta: Midsea Books.
- Avram, A. Andrei. 2012. “Some phonological changes in Maltese reflected in onomastics”. *Bucharest Working Papers in Linguistics* XIV (2): 99-119.
- Avram, A. Andrei. 2014. “The fate of the interdental fricatives in Maltese”. *Romano-Arabica* XIV: 19-32.
- Avram, A. Andrei. 2016. “Phonological Changes in Maltese: Evidence from Onomastics”. In *Shifts and Patterns in Maltese*, edited by Gilbert Puech and Benjamin Saade, 49-90. Berlin: De Gruyter Mouton.
- Avram, A. Andrei. 2017a. “Word-final obstruent devoicing in Maltese: Inherited, internal development or contact-induced?”. *Academic Journal of Modern Philology* 6: 23-37.
- Avram, A. Andrei. 2017b. “On the Phonology of Sicilian Arabic and of Early Maltese”. In *Advances in Maltese Linguistics*, edited by Mauro Tosco and Benjamin Saade, 3-37. Berlin: De Gruyter Mouton.
- Avram, A. Andrei. 2020. “The Diachrony of Word-Final Obstruent Devoicing in Maltese”. In *Maltese Linguistics on the Danube*, edited by Slavomir Cépľo and Jaroslav Drobný, 27-57. Berlin: De Gruyter Mouton.
- Barbera, Giuseppe. 1939. *Dizionario maltese-arabo-italiano. Con una grammatica comparata arabo-maltese*, vol. I, A – E. Beirut: Imprimerie Catholique.
- Borg, Alexander. 1975. “Maltese morphophonemics”. *Journal of Maltese Studies* 10: 11-28.
- Borg, Alexander. 1997. “Maltese Phonology”. In *Phonologies of Asia and Africa*, vol. 1, edited by Alan S. Kaye, 245-285. Winona Lake: Eisenbrauns.
- Borg, Alexander and Azzopardi-Alexander, Marie. 1997. *Maltese*. London and New York: Routledge.
- Burridge, Kathryn and Bergs, Alexander. 2017. *Understanding Language Change*. London and New York: Routledge.

- Bybee, Joan. 2016. *Language Change*. Cambridge: Cambridge University Press.
- Cachia, Lawrenz. 2000. *Habbew l-ilsien Malti*. Zabbar: Veritas Press.
- Cardona, Tony. 1997. *Introduzzjoni għal-lingwistika Maltija* [Introduction to Maltese Linguistics]. L-Msida: Mireva Publications.
- Cassola, Arnold. 1987-1988. “Una edizione diversa della lista di voci maltesi del seicento di Hieronymus Megiser”. *Journal of Maltese Studies* 17–18: 72–86.
- Chen, Matthew Y., and Wang, William S.-Y. 1975. “Sound change: Actuation and implementation”. *Language* 51: 255-281.
- Cohen, David. 1966. “Le système phonologique du maltais. Aspects synchroniques et diachroniques”. *Journal of Maltese Studies* 3: 1-26.
- Cohen, David. 1967. “Contribution à la phonologie diachronique du maltais”. In *Verhandlungen des zweiten internationalen Dialektologenkongresses*, edited by Ludwig E. Schmitt, 164-171. Wiesbaden: Steiner.
- Cohen, David and Vanhove, Martine. 1991. “La *Cantilène* maltaise du 15ème siècle: remarques linguistiques”. *Comptes-rendus du Groupe linguistique d'études chamito-sémitiques* XXIX–XXX: 177-200.
- Cowan, William. 1964. “An early Maltese wordlist”. *Journal of Maltese Studies* 2: 217-225.
- Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*, sixth edition. Oxford: Blackwell Publishing.
- Cser, András. 2015. “Basic Types of Phonological Change”. In *The Oxford Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph Salmons, 193-204. Oxford: Oxford University Press.
- Dresher, B. Elan. 2015. “Rule-Based Generative Historical Phonology”. In *The Oxford Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph Salmons, 501-521. Oxford: Oxford University Press.
- Fiorini, Stanley. 1987–1988. “Sicilian connexions of some medieval Maltese surnames”. *Journal of Maltese Studies* 17–18: 104-138.
- Hull, Geoffrey. 1994. “Review of Arnold Cassola.” *The Biblioteca Vallicelliana: Regole per la lingua maltese. The earliest Extant Grammar and Dictionary of the Maltese Language*. *Journal of Semitic Studies* XXXIX (2): 391-395.
- King, Robert D. 1969. *Historical Linguistics and Generative Grammar*. Englewood Cliffs, NJ: Prentice Hall.
- McMahon, April M. S. 1994. *Understanding Language Change*. Cambridge: Cambridge University Press.



- Phillips, Betty S. 2015. "Lexical Diffusion in Historical Phonology". In *The Oxford Handbook of Historical Phonology*, edited by Patrick Honeybone and Joseph Salmons, 359-373. Oxford: Oxford University Press.
- Uguzzoni, Arianna. 1977. *La fonologia*. Bologna: Zanichelli.
- Vanhove, Martine. 1994. "La langue maltaise un carrefour linguistique". *Revue du monde musulman et de la Méditerranée* 71 (1): 167-183.
- Vanhove, Martine. 2000. "Le maltais et les interférences linguistiques". In *Language in the Mediterranean Area. Typology and Convergence (II progetto MEDTYP: Studio dell'area linguistica mediterranea)*, edited by Sonia Cristofaro and Ignazio Putzu, 187-199. Milan: Franco Angeli.
- Vella, Olvin. 2013. "A Gozitan dialect for Standard Maltese in the eighteenth century". Paper presented at the 4<sup>th</sup> International Conference on Maltese Linguistics, 17-19 June 2013, "Lumière" University, Lyon.
- Wettinger, Godfrey. 1968. "The distribution of surnames in Malta in 1419 and in the 1480s". *Journal of Maltese Studies* 5: 25-48.
- Wettinger, Godfrey. 1983. "Some grammatical characteristics of the place-names of Malta and Gozo in early modern times". *Journal of Maltese Studies* 15: 31-68.

## Notes

- 
- <sup>1</sup> The <th> also reflects word-final obstruent devoicing.
- <sup>2</sup> Abela (1647: 73) does not provide the translation.
- <sup>3</sup> Not translated in Abela (1647: 16).
- <sup>4</sup> Where <ح> represents [ħ].
- <sup>5</sup> Where <ه> represents [h].
- <sup>6</sup> Where <ع> represents [ʕ].
- <sup>7</sup> Where <غ> represents [ɣ].
- <sup>8</sup> Where <ق> represents [q].
- <sup>9</sup> Where <چ> represents [tʃ].
- <sup>10</sup> The etymon proposed by Barbera (1939: 237) is more likely than Italian *borsa*, given in Aquilina (1987a: 138).
- <sup>11</sup> It is more plausible to assume a Sicilian origin of the Maltese word rather than deriving it etymologically from Italian *pericoloso*, as in Aquilina (1987b: 1049)
- <sup>12</sup> Aquilina (1987b: 1405).
- <sup>13</sup> According to the orthographic conventions of Catalan, <tx> represents [tʃ].
- <sup>14</sup> The digraph <ph> presumably stands for [p].
- <sup>15</sup> Aquilina (1987a: 175) only proposes the Italian etymon..
- <sup>16</sup> Aquilina (1987b: 1336).
- <sup>17</sup> Aquilina (1987b: 1376).
- <sup>18</sup> The trigraph <sch> represents [ʃ].
- <sup>19</sup> The digraph <tc> represents [tʃ].

---

<sup>20</sup> Aquilina (1987a: 1332).

<sup>21</sup> For a comparison of the phonology of Sicilian Arabic and early Maltese see Avram (2017b).

<sup>22</sup> Where <th> = *t*.

<sup>23</sup> 'Its sound is as in Latin and in Italian, yet sometimes it is pronounced with a lisp: as can be seen at the end of this book, where all the words pronounced with a lisp are found'. The list of words still containing interdental fricatives does not, however, appear at the end of Vassalli's (1796) dictionary.

<sup>24</sup> This grapheme designed by Vassalli stands for [ʃ].

<sup>25</sup> See also Borg (1975: 19-20).

<sup>26</sup> Except for some recent, phonologically non-integrated borrowings from English (Ray Fabri, University of Malta, p.c. June 2017).

## CHAPTER ELEVEN

# ENGLISH POSITIONAL SCHWA: A SIMILAR CATEGORY TO ESTABLISH BY ROMANIANS

ELENA-RALUCA CONSTANTIN

The Speech Learning Model (henceforth SLM: Flege 1988, 1995) assumes that the learnability of an L2 sound is inversely correlated with its similarity to an L1 sound. It is easier to develop a new category in L2 for what is referred to as a new sound, i.e., not similar to any L1 sounds. Although the process of equivalence classification reduces the ability to form a new category for a similar sound, it does not block learnability. For such cases, a merged L1-L2 category will be used (Flege 2002). The new versus similar sound differentiation also has a bearing on the accuracy of L2 sound production. A new sound will ultimately be produced more accurately than the merged L1-L2 diaphone. The production of the L2 English schwa of one group of Romanian informants (L1 Romanian; L2 English) was looked into to test the predictions that SLM posits. The stimulus included a word-final schwa as in English this schwa has vowel consistency (Flemming and Johnson 2007). The results showed that, in accordance with SLM, the L1 Romanian group exhibited signs of equivalence classification with low L2 English schwa accuracy regardless of level of proficiency.

Keywords: schwa vowel, equivalence classification, phonological proficiency, similar category

### **1. Introduction and preliminary remarks**

The aim of the current paper is to explore phonetic learning and phonetic approximation of a similar L2 sound within the theoretical framework of Flege's (1995, 2002, 2005) Speech Learning Model (henceforth SLM). In this vein of thought, the layout of the paper is as follows: first of all, the

hypotheses for category formation of similar sounds that derive from the assumptions of the SLM are overviewed, hereby including the process of equivalence classification which predicts different learnability and accuracy for new and similar sounds. The outcomes which the process of equivalence classification entails for the second language sound as well as the native sound are subsequently explained followed by a description of how language experience impacts category assimilation and dissimilation. The assumptions of SLM are then applied to the English similar schwa acquired by Romanians. Finally, the design of the study is laid out and the findings discussed along with future tenets to be considered for further investigation.

Basically, this study investigates the acquisition of L2 English schwa in terms of the correspondence between the target and the L1 Romanian native segment. The main assumption with strong empirical foundation is that the degree of phonetic similarity correlates inversely with L2 phonemes' learnability. This is the basic tenet of SLM, which the present study is mainly grounded in. Given the hypotheses generated by SLM, L2 segmental acquisition is envisaged as a process dependent on the interaction between the native and the target phonetic systems. A basic assumption the model builds on is that the mechanism for phonetic category formation is not constrained by age, being applied effectively in L2 learning. The SLM also posits that the phonetic systems responsible for both perception and production remain flexible and on encountering L2 sounds they can be reorganized by adding new phonetic categories or modifying existing ones (Flege 1995: 233). Phonetic evidence for the adaptability of phonetic systems is presented and thoroughly discussed in research studies concerned with native-like oral speech production that can be achieved even by late learners (Flege and MacKay 2004).

The degree of similarity between L1 and L2 phonemes is essential for the L2 sounds' learnability, with the specification that the SLM maintains that the more distant an L2 sound from the closest L1 category is, the more learnable this sound is likely to be, i.e., the more likely it is that a new category for this sound will be established. The more phonetically similar an L2 sound is to an L1 category, the less likely it is that the learner will set up a relevant phonetic category. This is due to the operating cognitive mechanism of equivalence classification which hinders the formation of categories for similar sounds since they are perceived as realizations of existing L1 categories (Flege: 1987). Nevertheless, this mechanism impedes, but does not fully block the formation of new categories for similar sounds. Phonetic learning is feasible without category formation because cross-language subcategorical differences are auditorily accessible to language learners (MacKay et al. 2001:517). For speech sounds which

represent one and the same phonetic category (i.e., they are subcategories), a merged category, a composite that combines the properties of the corresponding L1 and L2 speech sounds will be set up (Flege: 2005).

To clarify the assimilation vs. dissimilation discrimination, the SLM proposes that when learners are unable to create a new category for an L2 vowel because it is too similar to an existing L1 vowel, the two vowels will eventually form a composite – coming to resemble one another (Flege: 2005).

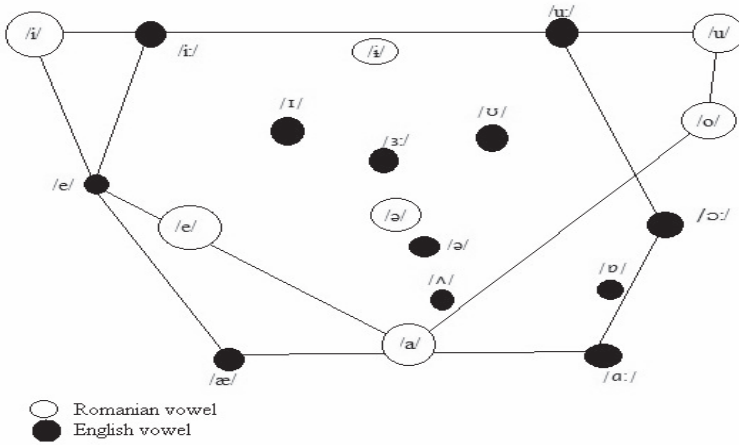
In order to verify how Flege's equivalence classification (1995) operates with respect to the acquisition of similar sounds (1995), the present study was guided by the following research questions that generate two main hypotheses:

- Q1 Do Romanians, as foreign language learners of English set up a new category for schwa as a similar sound they have to acquire?
- H1 I assume that although the mechanism of equivalence classification will reduce the Romanians' ability to form a new category for the similar sound, that is the target English schwa, it does not block the acquisition of the elicited target. In this particular case, a merged L1-L2 category will be used to produce the perceptually equated Romanian L1 and English L2 schwa vowel in line with Flege (2002).
- Q2 To what extent is the Romanians' accuracy of similar sound production influenced by level of proficiency?
- H2 As for the acquisition of schwa as a similar sound, in light of SLM, I assume less influence of level of proficiency because of the merged diaphone, with the Romanians' production of the L2 English schwa vowel being either linear or slightly different across the three levels of proficiency<sup>1</sup>.

### **1.1. The L1 Romanian schwa vowel: a similar sound**

In order to test the predictions of SLM with regard to similar L2 phone accuracy and the way one's level of proficiency affects both learnability and accuracy, I compared the production of the L2 English schwa of the L1 Romanian group with the native English reference provided by the literature (Chițoran 2002). In Romanian there is a schwa though with a slightly different place of articulation which classifies as a similar sound to the

English schwa. Figure 1 below illustrates that the Romanian schwa is lower and more towards the back than its English counterpart. Chițoran (2002) provides an accurate measurement of these articulatory differences, with the specification that the cited authors used the same measurement scales when establishing the articulatory points for the two schwas.



**Fig. 1: English and Romanian monophthongs. Synthetic chart.**

(Adapted from Chițoran et al 1984: 36)

Considering the Romanian and the English vowel plots in Figure 1 below, schwa stands for a case in point as the present experimental research is aimed at analyzing the production of similar phones in a foreign language, with a similar L2 phone differing systematically from an easily identifiable counterpart in LI. For example, /t/ is found in both French and English, but it is implemented as a short-lag stop with dental place of articulation in French, and as along-lag stop with alveolar place of articulation in English. The /u/ of French and English must also be classified as similar. For /u/ is realized with somewhat lighter and more variable second formant (F2) frequencies in English than French (Flege 1987: 48). Since in Romanian schwa stands for a monophthong which is similar to its English counterpart, I anticipate Romanians will not establish a distinct category to the English schwa.

## 2. The phonetics of English and Romanian schwa vowels

More often than not English schwa is read as a weak or reduced vowel because it is the mere result of neutralization of vowel quality contrasts (Flemming 2009). Due to vowel reduction or resistance to being stressed, schwa is also commonly restricted to unstressed syllables in English. As a matter of fact, the basis for the weakness of schwa has been the subject of much research by phonologists (Van Oostendorp 2000) who made a clear-cut distinction in their recent proposals between the nature of English schwa as a mid-central vowel, and English schwa as a vowel that lacks a well-defined target, and so assimilates strongly to surrounding segments, exhibiting substantial variation in its vowel quality. Thus, research conducted in the literature (Flemming 2009) indicate the existence of both kinds of schwa vowels: a true mid central vowel and a contextually-variable vowel, data which are in line with Lass (2007).

Furthermore, Flemming and Johnson (2007) acknowledged major phonetic differences between schwa vowels in word-final position, as in (1), and schwa vowels in other positions, as in (2).

- (1) china /tʃaɪnə/
- (2) suppose /sə'pəʊz/  
comma /kɒmə/  
probable /'prɒbəb(ə)l/

(Flemming and Johnson: 2007)

Whereas word-final schwa vowels have a relatively consistent vowel quality, usually mid central, word-internal schwa vowels are relatively high and vary contextually in backness and lip position. Therefore, variability of non-final schwa (particularly F2) was accounted for in terms of assimilation by context. Still, what makes this kind of schwa more variable than a full vowel? Flemming (2004) argues that two main related factors are involved: word-medial schwa is (i) very short, and (ii) does not minimally contrast with other vowel qualities. These two factors are correlated since it is the short duration of non-final unstressed syllables that favours the neutralization of vowel quality contrasts in these contexts.

“to realize a particular vowel quality in a word, it is necessary to move from the articulatory position of the previous segment to the target for the vowel and then on to the position of the following segment. As the duration of the vowel decreases, it can become difficult to complete the required movements, especially if the vowel target is far from the targets

for the preceding or following segments, because the articulators would have to move too fast to complete the movements in the time available.” (Flemming 2009).

As a matter of fact, the schwa vowel has an exceptional position, displaying a high level of context-dependency and a huge amount of variability. It is this high-context sensitivity in particular that has led to the assumption that schwa can also be targetless when not surfacing in unstressed word-final position. Therefore, schwa is phonetically realized with an active gesture that is, however, overlapped by the gesture of the following full vowel or a vowel underspecified for tongue position (Browman & Goldstein 1992: 26).

Basically, it is susceptibility/ unsusceptibility to coarticulation that makes English variable schwa in non-final word-position not be analyzed as a particular vowel quality, and the mid-central schwa found in word-final unstressed syllables be analyzed as a distinct category exhibiting a particular vowel quality.

In addition, it is crucial for the current research to distinguish two main functions of the English schwa: anaptyctic (3) on the one hand, and positional (4) on the other hand.”

- (3) today /tə'dei/  
 (4) support /sə'pɔ:(r)t/  
 (5) computer /kəm'pjʊ:tə(r)/ Heselwood (2007: 148)

In (5), the second occurrence of schwa is positional: the final vowel segment in *computer* (this word is followed here by a pause) functions, phonologically, as a distinctive segment (in opposition to the constitutive function of a segment) by virtue of the opposition segment position / empty position. The last schwa in (6) is also anaptyctic, since it is inserted into the sequence as a vowel sound, qualitatively assimilated to surrounding or adjacent sounds and for that reason different from theoretical canonical schwa, with the view to easing pronunciation.

- (6) thataway /ðætəweɪ/ Heselwood (2007: 148)

Anaptyctic schwa vowels are prone to be influenced by the phonetic context to a more considerable extent than positional schwa vowels. Thus, anaptyctic schwa vowels are employed as a mere type of epenthesis meant to facilitate the pronunciation of a consonant cluster, whereas positional schwas are used as phonologically distinct segments.



Similarly, the Romanian schwa vowel also has an anatyptic function beside the positional function, since targetless schwa in Romanian was also called in the literature “une voyelle vicaire/ sonus vicarius”. As Avram (1990: 9) claims in his research “en tenant compte de la "fonction" qui vient d'être mentionnée et aussi d'autres particularités des voyelles [ə] et [i], Sextil Puşcariu a nommé ces deux sons des *voyelles vicaires*”.

Let us consider the examples in (7a-c) below.

- (7) a. alt[ə]cine  
 b. opt[i]sprezece  
 c. ours[ə]blanc (Avram 1990: 9)

The Romanian central vowels in (7a), (7b) respectively, have the same function as their French equivalent, namely -"un rôle de lubrifiant" (Avram 1990: 9).

As Avram (1990: 9) points out, the insertion of a parasite vowel in consonant clusters that are difficult to utter is a frequent phonological phenomenon irrespective of the occurrence of the parasite vowel with respect to the morphemic boundaries

To avoid too much variation brought about by contextual coarticulation, our experiment is aimed at investigating the latter type of schwa, that is, positional schwa, a distinct segment on its own with inherent vowel consistency as proven insofar.

Unlike the English schwa, the Romanian schwa is consistent with both stressed and unstressed positions, with some differences in the second formant (henceforth F2) between the two occurrences, as shown in Table 1 that illustrates the mean formant values of the Romanian schwa, phonetically transcribed as /Λ/, for each gender, in each stress condition (Renwick 2012: 158). Basically, Table 1 encapsulates the values consistent with the first (F1) and second formants (F2) of the Romanian schwa vowel across male and female subjects, which shows that Romanian schwa has independent phonetic space from the other vowel phonemes that are associated with other acoustic measurements in terms of F1 and F2. There is a clear-cut distinction to be made between the values exhibited by schwa in unstressed syllables (referred to as UNS in Table 1) and the values that tally with schwa in stressed syllables (referred to as STR in Table 1).

			/a/	/e/	/i/	/ɨ/	/o/	/u/	/ɯ/
F	F1	STR	897	603	377	444	591	411	636
		UNS	856	552	333	450	573	406	583
	F2	STR	1463	2095	2720	1600	1003	1106	1503
		UNS	1473	1961	2745	1850	1073	1324	1595
M	F1	STR	679	495	317	381	497	363	519
		UNS	685	438	294	392	494	365	490
	F2	STR	1302	1737	2151	1482	993	1116	1377
		UNS	1239	1710	2149	1683	986	1232	1446

**Table 1. Mean standard deviations (Hz) for the first and second formants for vowel tokens pooled across speakers, separated by gender and stress condition (14 female speakers; 3 male speakers. Non-normalized data).**

(Renwick<sup>2</sup> 2012: 158)

One of the findings reported by Renwick (2014: 101) is that the Romanian schwa vowel occupies its own acoustic space, and does not exhibit the degree of variability of formant values that is expected from a targetless reduced vowel. The data analyzed in Renwick (2014: 101) obviously show that durationally, the Romanian L1 schwa vowel patterns with other vowels of its height class and thus is not reduced in dimension. For these reasons, the Romanian L1 schwa vowel should be treated unambiguously as a full vowel, even in unstressed position where it is contrastive as a morpheme (Anghelina 2008: 529).

Moreover, as Chițoran (2002: 210) points out the alternation of stressed [a] with unstressed schwa may be interpreted as an instance of vowel reduction, similar to that typical of English, for instance. Nevertheless, schwa does not stand for a phonologically reduced vowel in all contexts in Romanian, and it has phonemic status. It may surface under stress, and there is evidence of underlying schwa in various forms. The roots in (8a) for example, all feature underlying schwa, which surfaces in both stressed and unstressed positions. The schwa vowels in (8b) do not alternate with any other vowels, and their occurrence is unpredictable, thus supporting the view that they are part of the underlying representation of these forms. Let us consider the minimal pair in (8c) that brings further evidence in favour of the Romanian schwa vowel as an underlying phoneme.

- (8) a. mătură /mətʉrə/ “broom”– măturică /mətʉrikə/ “broomie”  
 (diminutive)  
 pătură /pətʉrə/ “blanket” – păturică /pətʉrikə/ “blankie”  
 (diminutive)

- b. pământ /pəmînt/ “earth”  
 părinte /pərînte/ “parent”  
 grătar /grətar/ “grill”
- c. păr /pər/ “hair”  
 par /par/ “pole” Chițoran (2002: 210)

Reduction alternations are frequent in Romanian (9), as Romanian exhibits phenomena of derived environment effects of stressless vowel reduction (Khanjian 2009:185).

As a matter of fact, in Romanian, [á] turns into [ʌ] when stress shifts (Steriade 2008):

- (9) a. /sʌrák/ “poor”      /sʌrʌk-úts/ (diminutive) “poor”  
 b. /papúk/ “slipper”      /papuţ-él/ (diminutive) “slipper”

Moreover, given the Romanian vowel space in Figure 2 (Sarlin 2014: 18), we could deduce that the Romanian schwa is backer when compared to the English positional schwa and this is a difference to consider while investigating the required tokens.

In a nutshell, considering schwa in phonological theory, there are two approaches that categorize the schwa vowel in the literature. First, there is a two-fold division of schwa varieties into a genuine mid-central schwa and a contextually variable schwa (Flemming and Johnson 2007). Second, there is a latter categorization positing three types of schwas (Oostendorp 2000, Veloso 2007, Constantin 2019). Such categorization covers three major types of schwa to be found both in English and Romanian: ‘e-schwa’ or epenthetic schwa - defined as schwa that results from epenthesis which usually alternates with zero, as shown in (6) and (7); ‘r-schwa’ or vowel-reduction schwa - which is schwa that alternates with a full vowel as the result of vowel reduction, as illustrated in (9a) for Romanian.

In English all unstressed vowels get reduced to schwa. Consider the example in (10) where the full unstressed vowel /u/ in the fourth syllable was reduced to schwa since the stop /t/ occurring in the coda requires a low sonority down step.

- (10) Connecticut /kə'netikət/ (Oostendorp 2000)

The third type of schwa’s-schwa’ or stable schwa stands for schwa present at the underlying representation. This kind of schwa is neither the result of epenthesis nor vowel reduction, as illustrated in (5) and (8c).

Among all types of schwas existing in both English and Romanian, it is mid-central schwa to be investigated with experimental data throughout the present study, given its lack of variation on the one hand, and the aims of my research, on the other hand. I think vowel-reduction issues, schwa epenthesis overviews and underlying representation analyses, which are addressed by the latter approach, exceed the scope of the current study since they generate a distinct research path to follow.

### **3. Methodology and experimental conditions**

#### **3.1 Participants' background**

In order to validate the predictions of SLM with regard to equivalence classification, a research group was needed. An aggregate of 33 Romanian respondents (5 females and 28 males aged 21) were selected for the oral production experiment, with Romanian as an L1 and English as an L2. Speakers of some other foreign languages, typically German, were also identified in the group, an aspect we considered negligible for the scope of our research. Three levels of proficiency were considered in English as an L2. The Romanian subjects tested belonged to three distinct categories according to their level of proficiency. Thus, 11 Romanian respondents had a B1 level of English, 13 Romanians were assigned a B2 level of English and the other 9 subjects a C1 command of English. It is to note that no participant within the research group reported impaired language or hearing.

All Romanian subjects were asked to fill in a thorough background questionnaire which elicited the participants' phonetic and linguistic background. More precisely, the informants were asked to provide personal data with respect to the number of foreign languages they spoke, the onset age and the end age for each target language. Added to all these, they were also required to mention the number of hours in the target language they benefited from both in and extra class. Other main issues such as access to institutionalized phonetic training and any long-term stays in the target language country were also included in the questionnaire.

It was compulsory for all Romanian informants to sit for an ERASMUS+ online language test since they all joined mobilities abroad in various receiving countries. The language test complied with the European Language Passport, which is a standardized template for self-assessment of language skills aimed at testing the non-natives' reading comprehension abilities, listening comprehension abilities, grammar and vocabulary skills as well as the speakers' performance with respect to key communicative phrases.

### 3.2 Oral production stimuli

All Romanian informants were required to read citation forms out loud, that is two-syllable words with word-final schwa preceded by a voiceless obstruent. This was the phonetic environment adopted so that any contextual variation or coarticulation would be avoided. A number of fillers were also inserted for distraction as the procedure goes. A word list of the recorded schwa vowel tokens appears in Table 2 below. The target vowel encapsulated in Table 2 is bold.

	Transcription	Orthography	Gloss	Stress
1.	/ˈsi: kə/	seeker	–	UNS
2.	/ˈfatə/	fată	girl	UNS
3.	/ˈti: tʃə/	teacher	–	UNS
4.	/ˈkapə/	capă	cloak	UNS
5.	/kəmˈpjʊ:tə/	computer	–	UNS
6.	/ˈfrikə/	frică	fear	UNS
7.	/ˈbɪtə/	bitter	–	UNS
8.	/ˈpatə/	pată	stain	UNS
9.	/ˈflæʃə/	flasher	–	UNS
10.	/ˈsəpə/	sapă	shovel	UNS
11.	/ˈkɪsə/	kisser	–	UNS
12.	/ˈmɑrfə/	marfă	goods	UNS

**Table 2. Word list featuring word-final schwa vowel tokens in English and Romanian.**

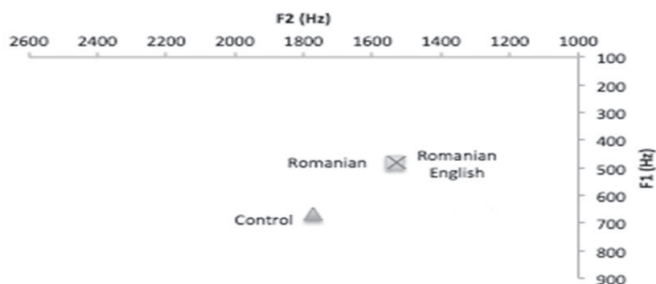
(Sypiańska and Constantin 2018)

### 3.3 Procedure

As for the recordings, they were carried out in a sound-treated booth within “Ferdinand I” Military Technical Academy of Bucharest, with all Romanian participants being explicitly instructed to avoid rhotacisation of the final vowel, which could have altered the values of the formants needed for the present study. The Romanians’ tendency would have been to utter a sequence of schwa+ the full phonetic realization of /r/. All rhotacised tokens that were subsequently identified in the speech samples were subsequently removed. Then F1 and F2 values were measured in Praat manually.

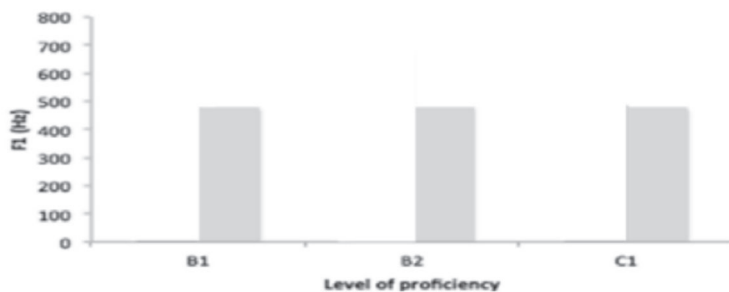
In the absence of a control group, the values in Renwick (2012: 158) were targeted. Since those values were obtained without normalization, I decided to keep the data collected from Romanians non-normalized.

#### 4. Findings and future tenets



**Fig. 2. Vowel chart comparing the production of the schwa by the Romanian research group with the English reference and the Romanian schwa as produced by the L1 Romanian group.**

Figure 2 illustrates the fact that the Romanian F1 and F2 values overlap the English targets, with Romanians being more oriented towards L1 in their schwa vowel production. This further means that the Romanian subjects did not succeed in setting up a phonetic new category for the English schwa vowel. Still, there are further issues to clarify when conducting future research on canonical schwa vowel production by Romanians in their capacity of non-native speakers of English as an L2. Basically, it was not proven by the current data whether there was complete assimilation of the elicited schwa vowel token by its Romanian counterpart, or whether there was an L1-L2 merge, with Romanians ultimately producing a composite of the two schwa vowels. With the view to bringing phonetic evidence to support one of the previously - mentioned phonological hypotheses, the design of the current methodology needs improvement. A control group of Romanians with no knowledge of English or any other foreign language should be selected for future research. This will definitely prove if there is negative transfer from Romanian L1 into English L2 or a mere L1-L2 composite in the native Romanians' production of the English schwa vowel.



**Fig. 3. Romanians' production of the English schwa vowel across three levels of proficiency**

Figure 3 shows that Romanians were constant in their English schwa vowel production irrespective of their level of proficiency. No improvement was recorded from the B1 category of Romanian subjects to the B2 class, and upwards to the C1 category.

For future research, for more precision in establishing Romanians' level of proficiency, foreign accent rating will be resorted to. In this vein of thought, Romanian phonetically-trained judges will be asked to assess the Romanians' production of the fillers and rate each word for global foreign accent on a 9-point scale ranging from 1 for native-like to 9 for heavy foreign accent.

## 5. Concluding remarks

Although there are some limitations to consider for future research as explained in section 4, the data collected for the current experimental study validate Flège's SLM (2005), proving that the degree of similarity between L1 and L2 phonemes is determining for the acquisition of L2 sounds. The phonetic realization of the L2 English schwa vowel was looked into since this central monophthong is a similar category in Romanian, even if backer than the English target Romanians had to acquire. Within the Romanian research group, it was the operating cognitive mechanism of equivalence classification which impeded the formation of a distinct category for the English schwa vowel since it was perceived as a mere realization of the existing L1 schwa vowel category. In this case, the results illustrate that the Romanian participants were not able to establish a new category for the English schwa since it was too similar to the existing Romanian vowel. In the future, a subsequent tenet is to be investigated. If Romanians were not able to set up a new category for the English schwa vowel, as proven

insofar, the appropriate operating mechanism is to be identified. There are two hypotheses the validity of which should be tested with further phonetic evidence to decide which one holds. Basically, if new category formation is not applicable, do the Romanian L1 and the English L2 schwa vowels merge and make up a Romanian-English composite, with Romanians being more L1-oriented in their schwa vowel production, or is there complete assimilation with the English target being fully assimilated by the Romanian existing category? Undoubtedly, this is a research path that needs further investigation.

Moreover, it was shown that the Romanian schwa vowel production of the English target was constant and linear irrespective of the respondents' level of proficiency with no differences recorded between B1, B2 and C1. Hence, the second research hypothesis was validated by the data under consideration since we assumed that the level of proficiency will not influence the Romanians' production of the schwa vowel because of the L1-L2 merged values. Still, foreign accent rating is needed to replace the placement test adopted for the current experiment with the view to assigning a more precise and appropriate level of proficiency to the Romanian informants.

There is one supplementary limitation to be considered when it comes to the current study: I think that in order to fully validate SLM a comparison within the same group of informants should have been made as to their production of an L2 vowel that is not present in L1 (or that is dissimilar to L1 vowels) when compared to their production of the L1 similar vowel from L2 schwa. One phenomenon that might contradict the SLM would be the acquisition of sounds such as /θ/ or /ð/, which do not have similar counterparts in Romanian but which are often approximated to /s/, /f/, /t/ or /z/, /v/, /d/, which, in turn, shows that speakers do not develop a new category. This issue is to be investigated in an upcoming paper.

## References

- Avram, Andrei. 1990. "Remarques sur les voyelles neutres du roumain et du français". *Word* 24: 8-13.
- Anghelina, Cătălin. 2008. "On Some of the Uses of Schwa (Ă) in Romanian Verbal Morphology". *Revue Roumaine de Linguistique*, LIII (4): 529-532.
- Browman, Carol, and Louis Goldstein. "Targetless" schwa: An articulatory analysis". In *Papers in Laboratory Phonology II*, edited by John Kingston and Mary E. Beckman, 26-56. Cambridge: Cambridge University Press.



- Chițoran, Dumitru, James E. Augerot, and Hortensia Pârlog. 1984. *The Sounds of English and Romanian*. Bucharest: Editura Universității din București.
- Chițoran, Ioana. 2002. *The Phonology of Romanian: A Constraint-Based Approach*. New York: Mouton de Gruyter.
- Conseil of Europe. Assessment grid. Retrieved from <https://rm.coe.int/CoERMPublicCommonSearchServices> on January 15<sup>th</sup> 2019.
- Constantin Elena-Raluca. 2019. "S-schwa production: A Case Study on Romanians". Paper presented at 6<sup>th</sup> Colloquium of Language Acquisition, University of Bucharest, November, 16 2019.
- Flege, James Emil. 1987. "The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification". *Journal of Phonetics* 15: 47-65.
- Flege, James Emil. 1988. "The production and perception of foreign languages". In *Human Communication and Its Disorders*, edited by Harris Winitz, 224-401. Norwood, NJ: Ablex.
- Flege, James Emil. 1995. "Second language speech learning: Theory, findings, and problems". In *Speech Perception and Linguistic Experience: Issues in Cross-Linguistic Research*, edited by Winifred Strange, 233-276. Timonium, MD: York Press, 1995.
- Flege, James Emil. 2002. "Interactions between the native and second-language phonetic systems". In *An Integrated View of Language Development: Papers in Honor of Henning*, edited by Petra Burmeister, Thorsten Piske, and Andreas Rohde, 217-244. Trier: Wissenschaftlicher.
- Flege, James Emil. 2005. "Origins and development of the Speech Learning Model". Keynote lecture presented at the 1st ASA Workshop on L2 Speech Learning Simon Fraser Univ., Vancouver, BC April 14-15.
- Flege, James Emil, and Ian MacKay. 2004. "Perceiving vowels in a second language". *Studies in Second Language Acquisition* 26 (1): 1-34.
- Flemming, Edward. 2004. "Contrast and perceptual distinctiveness". In *Phonetically-Based Phonology*, edited by Bruce Hayes, Robert Kirchner, and Donca Steriade, 232-276. Cambridge: Cambridge University Press.
- Flemming, Edward. 2009. "The Phonetics of Schwa Vowels". In *Phonological Weakness in English – From Old to Present-Day English*, edited by Donka Minova, 78-95. Hampshire: Palgrave MacMillan.
- Flemming, Edward, and Stephanie Johnson. 2007. "Rosa's roses: reduced vowels in American English". *Journal of the International Phonetic Association* 37 (1): 83-96.

- Heselwood, Barry. 2007. "Schwa and the phonotactics of RP English". *Transactions of the Philological Society*. 105 (2): 148-187.
- Khanjian, Hrayr. 2009. "Stress Dependent Vowel Reduction". *Berkeley Linguistics Society* 35 (1): 178-189.
- Lass, Robert. 2007. "On schwa: synchronic prelude and historical fugue". In *Phonological weakness in English: From Old to Present-Day English*, edited by Donka Minova, 47-77. Hampshire: Palgrave Macmillan.
- MacKay, Ian, Emil Flege J., Piske Thorsten, and Carlo Schirru. 2001. "Category restructuring during second-language speech acquisition". *The Journal of the Acoustical Society of America* 110 (1): 516-528.
- van Oostendorp, Marc. 2000. *Phonological Projection; A Theory of Schwa in Optimality Theory*. Berlin: Mouton-De Gruyter.
- Renwick, Margaret Elspeth Lambert. 2012. "Vowels of Romanian: historical, phonological and phonetic studies." PhD diss., Cornell University.
- Renwick, Margaret Elspeth Lambert. 2014. *The Phonetics and Phonology of Contrast*. Berlin: Mouton De Gruyter.
- Sarlin, Mika. 2014. *Romanian grammar* (2<sup>nd</sup> Edition). Helsinki: Books on Demand GmbH.
- Steriade, Donca. 2008. "Contour correspondence: The segmental evidence". Paper presented at the 11th International Symposium on Chinese Languages and Linguistics, National ChiaoTung University, Taiwan, May, 23-25.
- Sypiańska, Jolanta, and Elena-Raluca Constantin. 2018. "Phonetic learnability and accuracy of new and similar L2 sounds". Paper presented at the 12<sup>th</sup> International Conference on Native and Non-native Accents of English, University of Łódź.
- Veloso, João. 2007. "Schwa in European Portuguese: The Phonological Status of [ɨ]". In *Actes des/Proceedings of JEL '2007. Schwa(s). 5.èmes Journées d'Etudes Linguistiques*, edited by Olivier Crouzet and Jean-Pierre Angoujard, 55-60. Nantes: Université de Nantes.

## Notes

<sup>1</sup> The European passport of languages has been used as a system of reference here (the three levels being A, B and C in line with the assessment grid at <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?document>).

<sup>2</sup> In line with Steriade (2008), Renwick (2014: 13) takes a position on the standard transcription of the schwa vowel /ə/. In her studies, the schwa vowel is transcribed

---

/ʌ/. The latter transcription is substituted with the one in Table 1 as the usage of the standard phonetic transcription may be misleading, suggesting that the schwa vowel in question is a reduced one, which perhaps participates in phonological neutralizations in prosodically weak positions. According to Renwick (2014: 13), while the mid central vowel historically developed in unstressed syllables and was likely a reduced vowel, it functions synchronically as a full vowel and should be distinguished from reduced vowels.

## CHAPTER TWELVE

### SOME NOTES ON THE SYNTAX OF CORRECTIVE ‘CI’/ ‘BUT’ IN ROMANIAN

MIHAELA TĂNASE-DOGARU

The present paper aims at showing that Romanian possesses two types of corrective sentences. Following McCawley (1991), Vicente (2010), Toosarvandani (2013), these two types will be called anchored correctives and basic correctives. Anchored corrective sentences will be shown to consist of a coordination of clausal constituents while basic corrective sentences will be shown to coordinate sub-clausal constituents. Finally, the paper takes the first steps to suggest an alternative analysis of the syntax of corrective sentences based on multidominance.

Keywords: corrective ‘but’, anchored correctives, basic correctives, clausal coordination, sub-clausal coordination

#### 1. Introduction

This paper is concerned with the corrective use of ‘ci’ (‘but’) in Romanian and it addresses the issue of the kind of constituents ‘ci’ (‘but’) combines. Previous work in the domain of corrective adversative coordination assumed that corrective ‘but’ combines either clausal constituents (Vicente 2010) or sub-clausal constituents (Toosarvandani 2013, Franco 2016). Therefore, (1) - called ‘anchored form’ in McCawley 1991) - is analyzed as a combination of clausal constituents, while (2) - called ‘basic form’ in McCawley 1991 - is analyzed as a combination of sub-clausal constituents.

- (1) Gabriel didn’t drink milk, but coffee.
- (2) Gabriel drank not milk but coffee.

In Vicente (2010), it is shown convincingly that the syntax of corrective ‘but’-sentences involves clausal coordination, followed by ellipsis in the case of anchored form sentences (Vicente 2010).

- (3) Maria nu bea ceai ci cafea anchored corrective  
 Maria not drinks tea but coffee  
 “Maria doesn’t drink tea but coffee.”

Equally convincing is the demonstration in Toosarvandani (2013, 2014) and Franco (2016), according to which the syntax of corrective ‘but’-sentences involves sub-clausal coordination of DPs, PPs, AdvPs, in the case of basic form sentences.

- (4) Maria bea nu ceai ci cafea basic corrective  
 Maria drinks not tea but coffee  
 “Maria doesn’t drink tea but coffee.”

The present paper analyzes corrective ‘but’-coordination in Romanian, in an attempt at suggesting a unifying approach to the syntax of both anchored correctives and basic forms. In so doing, it will make use of available theories on multidominance (de Vries 2005, 2009, 2013, Citko 2012, van Riemsdijk, 2000, 2001, 2006), which will prove instrumental in dispensing with the need to posit different syntactic structures for anchored and basic correctives.

## 2. The data

In Romanian there are three versions of ‘but’: counterexpectational ‘but’, semantic opposition ‘but’ and corrective ‘but’.

The counterexpectational ‘but’ (‘dar’ in Romanian) (Anscombe and Ducrot 1977, Vicente 2010, Toosarvandani 2013, Franco 2016), illustrated in (5), introduces the implicature that the second conjunct is unexpected given the first conjunct (Vicente 2010). In (5), the first conjunct introduces the idea that there is a set of assumptions about strict teachers which does not include being approachable, a premise which is contradicted by the assertion of the second conjunct.

- (5) Profesorul e strict, dar abordabil.  
 Teacher.DEF is strict, but approachable.  
 “The teacher is strict, but approachable.”

The semantic opposition ‘but’ (‘dar’ / ‘iar’ in Romanian) (Toosarvandani 2013, 2014, Zafiu 2005) simply introduces the conjunction of two propositions, neither of which is denied, but which contain an opposition in terms of a relevant dimension (height, in this case), as illustrated in (6).

- (6) George e înalt, dar / iar Dan e scund.  
George is tall, but Dan is short.

The corrective ‘but’ (‘ci’ in Romanian) (Vicente 2010, Toosarvandani 2013, Franco 2016) is used to deny the proposition expressed by the first conjunct and introduces a closely related but true proposition (Vicente 2010). As illustrated in (7), the corrective reading is created by the combination of the denial of the first conjunct and the assertion of the second conjunct.

- (7) Gabriel nu bea lapte ci ceai.  
Gabriel not drinks milk but tea.  
“Gabriel does not drink milk but tea.”

Therefore, the difference between the counterexpectational and the corrective ‘but’ is that, with corrective ‘but’, what is expressed in the first conjunct is not true, while the second conjunct is true under the same circumstances (Steindl 2013, Franco 2016).

The paper addresses the issue of the kind of constituents corrective ‘but’ (‘ci’ in Romanian) combines: clausal constituents (Vicente 2010) (8) or sub-clausal constituents (9) (Toosarvandani 2013, Franco 2016):

- (8) Gabriel nu a băut lapte, ci cafea anchored corrective  
Gabriel not has drunk milk, but coffee  
“Gabriel didn’t drink milk, but coffee.”
- (9) Gabriel a băut nu lapte ci cafea basic corrective  
Gabriel has drunk not milk but coffee  
“Gabriel drank not milk but coffee.”

The section has looked at the different uses of ‘but’ in Romanian: the counterexpectational, the semantic opposition and the corrective ‘but’, respectively, for which Romanian possesses a different lexeme, ‘ci’. The next section looks at some features of corrective ‘but’ in a variety of languages.

## 2.1 Some features of corrective ‘but’

With the exception of Toosarvandani (2014), who claims that ‘but’ is polysemous in the manner of modal verbs, it is usually assumed in the literature, following Anscombe and Ducrot (1977), that, in English, there are two different homophonous lexical items ‘but’, one with a counterexpectational meaning and the other with a corrective meaning:

- (10) The girl is tall but no good at basketball. counterexpectational  
 (11) Gabriel didn’t drink beer but champagne. corrective  
 (Vicente 2010)

This is based on the fact that, in many languages, there are distinct lexical items for the counterexpectational and the corrective use:

- (12) a. Susana es pobre **pero** onesta.  
 Susana is poor but honest. Spanish, counterexpectational
- b. Gabriel no bebió cerveza **sino** champán.  
 Gabriel not drank beer but champagne.  
 “Gabriel didn’t drink beer but champagne.” corrective  
 (Vicente 2010: 386ff)
- (13) a. Der Raum ist klein, **aber** mein.  
 “The room is small, but mine.”  
 German, counterexpectational
- b. Peter ist nicht dumm, **sondern** faul.  
 “Peter is not stupid but lazy.” corrective  
 (Abraham 1979: 90ff)
- (14) a. Hu lo kalkelan **ela** ish asakim.  
 He no economist but man business.  
 “He is not an economist but a businessman.”  
 Hebrew, counterexpectational
- b. Hu lo kalkelan, **aval** hu ish asakim.  
 He no economist but he man business  
 “He is not an economist, but he is a businessman.”  
 corrective  
 (Dascal & Katriel 1977: 144)

- (15) a. Majid farānsavi ne-midune **balke** ālmāni midune.  
 Majid French not-knows but German knows  
 “Majid does not know French; he knows German.”  
 Persian, counterexpectational
- b. Majid farānsavi ne-midune **vali/amma** ālmāni midune.  
 Majid French not-knows but German knows  
 “Majid does not know French; though he should not know  
 German, he does.” corrective
- (Toosarvandani 2013: 26ff)

One of the most important characteristics of corrective ‘but’ (‘ci’ in Romanian) is the fact that it is only licensed by negation in the first conjunct (see also Vicente 2010, Toosarvandani 2013 a.o.). Negation in the first conjunct licensing corrective ‘but’ is metalinguistic negation, defined as ‘a device for objecting to a previous utterance’ (Horn 1989: 363ff), which entails that corrective ‘but’ is used to deny whatever is asserted by the first conjunct. This has two consequences for the syntax of corrective sentences.

First, since metalinguistic negation cannot be morphologically incorporated into word morphology, negative prefixes cannot license corrective ‘but’ (Vicente 2010: 384, Bosque 1980: 137):

- (16) a. \*This is improbable, but merely possible.  
 b. \*Esto es improbable, sino meramente posible.  
 “This is improbable, but merely possible.”  
 c. \*Acest lucru este improbabil, ci doar posibil.  
 This thing is improbable, but merely possible.
- (17) a. This is not probable, but merely possible.  
 b. Esto no es probable, sino meramente posible.  
 This not is probable, but merely possible.  
 c. Acest lucru nu este probabil, ci doar posibil.  
 This thing not is probable, but merely possible.

Secondly, since metalinguistic negation cannot license negative polarity items (Horn 1989, van der Wouden 1997: 69), clauses with corrective ‘but’ do not license negative polarity items in English or postverbal N-words in Spanish (Vicente 2010: 384):



- (18) a. I haven't (\*ever) been to Mexico but to Canada.  
 b. No he estado (\*nunca) en México sino en Canadá.  
 Not have been ever in Mexico but in Canada.  
 c. Nu am fost (\*niciodată) în Mexic ci în Canada.  
 Not have been (\*ever) in Mexic but in Canada.  
 "I have never been to Mexico, but to Canada".

In contrast, counterexpectational 'but' does not need sentential negation in the first conjunct; when negation is present, it licenses polarity items and postverbal N-words:

- (19) a. I haven't ever been to Mexico, but I have been to Canada.  
 b. No he estado nunca en México, pero he estado en Canadá.  
 Not have been ever in Mexico but have been in Canada.  
 c. Nu am fost niciodată în Mexic, dar am fost în Canada.  
 Not have been ever in Mexico but have been in Canada.

Having looked at a number of features corrective sentences exhibit, the next section focuses on the syntactic analysis of anchored correctives.

### 3. The Clause-Only Coordination Hypothesis (Vicente 2010)

Vicente (2010) analyzes instances of corrective 'but' with sentential negation, i.e., the anchored form (20) and shows that the difference between corrective and counterexpectational 'but' is syntactic, in that corrective 'but' always requires its conjuncts to be full clauses. He, therefore, postulates (21).

- (20) Gabriel no bebió cerveza **sino** champán.  
 Gabriel not drank beer but champagne.  
 "Gabriel didn't drink beer but champagne."

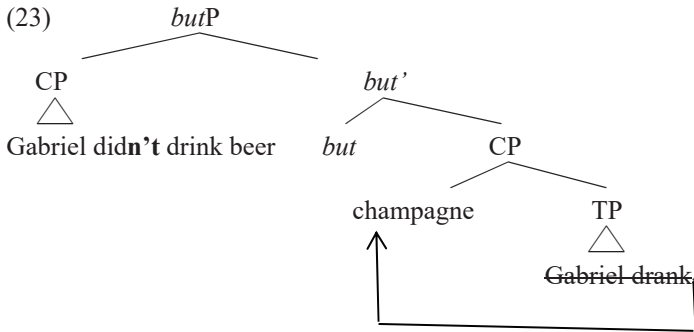
- (21) The syntax of adversative coordination  
 Corrective 'but' always requires its conjunct to be full clauses  
 Counterexpectational 'but' allows its conjuncts to be smaller than clauses. (Vicente 2010: 385)

Vicente (2010) lists various pieces of evidence in favor of this analysis.

First, the scope of negation is restricted to the first conjunct only, unlike cases of ‘and’-coordination, where the scope of negation may extend to the second conjunct:

- (22) Gabriel didn’t drink beer and champagne.  
 Gabriel nu a băut bere și șampanie.  
 Gabriel not has drunk beer and champagne

This is taken as an indication of the fact that negation is embedded in the first conjunct and it cannot scope over the second conjunct because of a lack of c-command.



Secondly, the impossibility of preverbal subject coordination speaks in favor of the combination of clausal constituents. Corrective ‘but’ cannot coordinate two preverbal subjects, in sharp contrast to ‘and’-coordination.

- (24) \*Two mathematicians but seven astrophysicists didn’t get their papers published.  
 \*Doi studenți ci trei profesori nu au plecat.  
 Two students but three professors not have left.

Thirdly, the impossibility of coordinating preverbal subjects with ‘but’ is explained by the Backward Anaphora Constraint (Ross 1967, 1969, Langacker 1969), which prohibits backward ellipsis within coordinate structures. In Spanish and Romanian, it is possible to coordinate right-peripheral subjects, which is predicted by a conjunction reduction analysis, where the second conjunct is part of an elided clause.

- (25) Nu au plecat doi studenți ci trei profesori [~~au plecat~~].  
 Not have left two students but three professors.  
 “Two students didn’t leave but three professors did.”

The fourth piece of evidence comes from the domain of attributive adjective coordination. Corrective ‘but’ cannot coordinate attributive adjectives, as shown in (26); if we assume a clausal coordination analysis, this would have to involve a combination of backward and forward ellipsis:

- (26) \*I didn’t read a short but long book.  
 [I didn’t read a ~~short~~ book] but [I ~~read a~~ long book]  
 I didn’t read a short book but a long one.  
 I didn’t read a short book, but [a long one [I ~~read~~]]

In Spanish and Romanian, since attributive adjectives are postnominal, it is possible to create the same effect by adding a PP to the right of the adjective:

- (27) \*Mauricio no ha leído un libro corto sino largo de Neal Stephenson  
 Mauricio not has read a book short but long by Neal Stephenson.  
 “Mauricio hasn’t read a short book by Neal Stephenson, but he has read a short one”.

Spanish, (Vicente 2010: 390)

- (28) \*Gabriel nu a jucat un joc scurt ci lung la PS4. Romanian  
 Gabriel not has played a game short but long at PS4.  
 “Gabriel hasn’t played a short game at his PS4 but he has played a long one”.  
 [Gabriel nu a jucat un joc scurt ~~la PS4~~] ci [Gabriel a jucat un joc lung la PS4].

The fifth piece of evidence comes from the domain of agreement phenomena. Starting from the assumption that, with coordination of clause-final subjects, the second conjunct is part of an elided clause, it is possible to formulate the following prediction: if the second conjunct belongs to a separate clause, it will not trigger agreement on the first conjunct verb and, therefore, a first conjunct effect will arise:

- (29) No se prezentó / \*presentaron un pianista  
 Not REFL showed.up.3SG / \*showed.up.3PL a pianist  
 sino tres trombonistas.  
 but three trombone players  
 “A pianist didn’t show up but three trombone players did.”  
 (Vicente 2010: 392)
- (30) Nu a /\*au plecat un student ci trei profesori.  
 Not have.3SG / \*have.3PL left a student but three professors.  
 “A student has not left but three professors have”.

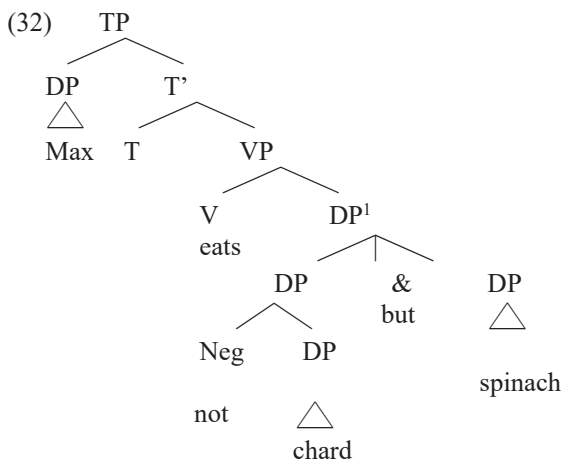
All the facts summarized above seem to point to an analysis of corrective ‘but’ in terms of clause-level coordination, followed by ellipsis. In what follows, the paper will focus on basic forms involving corrective ‘but’ coordination, with a view to showing that this type of corrective coordination involves coordination of sub-clausal constituents.

#### **4. The Sub-Clausal Coordination Hypothesis (Toosarvandani 2013)**

The clause-only coordination hypothesis (as Toosarvandani 2013 calls the hypothesis in Vicente 2010) cannot be maintained in view of cases such as (31b), where corrective ‘but’ coordinates two DPs, called basic form in McCawley 1991:

- (31) a. Max doesn’t eat chard but spinach  
           anchored form  
       b. Max eats not chard but spinach.  
           basic form  
(Toosarvandani 2013: 4)

While the anchored form (in 31a) is inarguably coordination of clausal constituents, in (31b) corrective ‘but’ conjoins two DPs, without ellipsis being necessary:



According to Toosarvandani (2013), the main piece of evidence in favor of analyzing basic forms of ‘corrective ‘but’ coordinations as sub-clausal coordinations comes from the fact that they behave like a single constituent. Various phenomena such as: fronting (33), it-clefting (34), and pseudoclefting (35) illustrate this point (all English examples from Toosarvandani 2013: 5ff).

- (33) Not Hermia but Helena now I love.  
 Nu studentul ci profesorul a plecat.  
 Not student.DEF but professor.DEF has left.  
 “Not the student but the teacher left.”
- (34) It’s not wordplay but weaponplay that’s needed.
- (35) What Wilde’s bon mot aroused was not Proust’s indignation but his compassion.  
 Ce a trezit interesul studentului a fost nu coperta  
 What awakened interest.DEF student.DEF.GEN was not cover.DEF  
 ci conținutul.  
 but content.DEF  
 “What awakened the student’s interest was not the cover but the content.”

Toosarvandani (2013) shows that corrective ‘but’ may coordinate a variety of sub-clausal constituents, with constituent negation adjoining to the first coordinate. Since, in the basic form, corrective ‘but’ need not

appear in sentence-final position, not all basic forms can be derived by clausal coordination followed by ellipsis.

(36) DP coordination

Not the neighborhood library perhaps, but the special engineering library will play a crucial role DP coordination

Nu conținutul ci coperta l-a  
 Not content.DEF but cover.DEF CL.3.SG.ACC.MASC-has  
 fermecat pe student.  
 enchanted DOM student  
 “It wasn’t the content but the cover that fascinated the student.”

(37) PP coordination

But the putting-forward, not of detailed and scientifically ‘finished’ hypotheses, but of schemata for hypotheses, has long been a function of philosophy.

A scris nu cu stiloul ci cu creionul.  
 (she)wrote not with pen.DEF but with pencil.DEF  
 “She didn’t write with her pen but with her pencil.”

(38) AdvP coordination

We were being not slowly but quickly poisoned  
 Mergeau nu încet ci repede.  
 walk.IMPERF.3PL not slowly but quickly  
 “They were walking not slowly but quickly.”

A further piece of evidence listed by Toosarvandani (2013) concerns islandhood. While, in the anchored form, the remnant cannot originate inside an island, as the examples (39–42) show, the basic form does not show island sensitivity, which is illustrated in (43–46).

(39) The Coordinate Structure Constraint (also discussed in Vicente 2010)

\*Alfonse didn’t cook rice and beans, but potatoes.  
 ??George nu a băut vin și bere, ci whisky.  
 George not has drunk wine and beer but whisky.  
 “George didn’t drink wine and beer, but whisky.”

- (40) Adjunct island (also discussed in Vicente 2010)  
 \*Jasper didn't choke when he saw Sally, but John.  
 ?? George nu s-a împiedicat când a văzut-o  
 George not REFL-has stumbled when has seen-CL.ACC.FEM  
 pe Ioana ci pe Maria.  
 DOM Ioana but DOM Maria  
 "George didn't stumble when he saw Ioana but when he saw Maria."
- (41) The subject constraint  
 \*That Alfonse ate the rice isn't fantastic but the beans.  
 \*Că George a vizitat muzeul nu e impresionant,  
 That George has visited museum.DEF not is impressive  
 ci castelul  
 but castle.DEF  
 "That George visited the museum is not impressive but that he visited the castle is."
- (42) The Complex NP Constraint  
 \*Alfonse didn't smash the vase that Sonya had brought from China, but from Japan  
 George nu a mâncat pastele pe care i  
 George not has eaten pasta.DEF.PL DOM which CL.DAT  
 le-a adus Maria de acasă ci de la restaurant.  
 CL.ACC.PL-has brought Maria from home but from at restaurant  
 "George didn't eat the pasta which Maria brought from home but he ate the pasta brought from the restaurant."
- (43) Alfonse cooked rice and not beans but potatoes.  
 George a băut bere și nu whisky ci vin.  
 "George has drunk beer and not whisky but wine."
- (44) Jasper choked when he saw not Sally but John.  
 George s-a împiedicat când a văzut-o  
 George REFL-has stumbled when saw-CL.ACC.FEM  
 nu pe Maria ci pe Ioana.  
 not DOM Maria but DOM Ioana.  
 "George stumbled when he saw not Maria but Ioana."

- (45) That Alfonse ate not the rice but the beans is fantastic.  
 Că George a vizitat nu muzeul ci castelul  
 That George has visited not museum.DEF but castle.DEF  
 e impresionant.  
 is impressive.  
 “That George visited not the museum but the castle is impressive.”
- (46) Alfonse smashed the vase that Sonya had brought not from China  
 but from Japan.  
 George a mâncat pastele pe care i le-  
 George has eaten pasta.DEF.PL DOM which CL.DAT. CL.ACC.PL-  
 a adus Maria nu de acasă ci de la restaurant.  
 has brought Maria not from home but from restaurant  
 “George ate the pasta which Maria brought not from home but  
 from the restaurant.”

Finally, it can be shown that basic correctives involving DPs can be parsed as subjects (47):

- (47) Nu un profesor ci un student a descoperit greșeala.  
 Not a teacher but a student has discovered mistake.DEF  
 “Not a teacher but a student discovered the mistake.”

The sequence ‘not a teacher but a student / nu un profesor ci un student’ is assumed in Bianchi and Zamparelli (2004) and Vicente (2010) to occupy a left-peripheral position, being derived by clausal coordination and ellipsis, while coordination is assumed to apply at the level of Focus projections in the left periphery, leading to the characterization of these sentences as *adjacent initial-edge coordinations*.

Following Rizzi’s (1997) demonstration of the fact that clauses can contain only one Focus Projection, Toosarvandani (2013) shows that the sequence ‘not a teacher but a student / nu un profesor ci un student’ does not occupy a left-peripheral position, since the left periphery can be occupied by a focused element (48):

- (48) THE NEUTRON, not a mathematician but a physicist discovered.  
 (Toosarvandani 2013: 839)
- (49) GREȘEALA, nu un profesor ci un student a  
 Mistake.DEF, not a teacher but a student has  
 descoperit-o.  
 discovered-CL.ACC.FEM  
 “The mistake, not a teacher but a student discovered.”



Therefore, the basic corrective ‘not a teacher but a student’ is derived by sub-clausal coordination and the two DP coordinates are merged in SpecT as canonical subjects.

The close investigation of the data shows that, in Romanian, the basic form does not involve clausal coordination followed by ellipsis, but sub-clausal coordination of DPs, PPs, AdvPs.

### 5. A note on agreement with corrective ‘but’

This section aims at showing that agreement patterns with anchored and basic correctives in Romanian reflect the difference in their syntactic structures, i.e., clausal vs. subclausal coordination.

In Vicente (2010), agreement facts clearly point to an analysis of anchored correctives as involving clausal coordination followed by ellipsis (50). When corrective *sino* links two clause-final subjects, a first conjunct effect arises:

- (50) No cometió / \*cometieron un error un pianista sino  
 Not made.3SG/ made.3PL a mistake a pianist but  
 tres trombonistas.  
 three trombone players.  
 “A pianist didn’t make a mistake but three trombone players did.”  
 (Vicente 2010: 392)

Basic correctives show a different pattern, as illustrated in (51):

- (51) Cometieron un error no un pianista sino tres trombonistas.  
 Made.3PL an error not a pianist but three trombone players  
 “A pianist didn’t make a mistake but three trombone players did.”

The examples in (52-55) show that the same pattern is valid in Romanian (see Franco 2016 for Italian):

- (52) Nu pleca / \*plecau un student ci  
 Not leave.IMPERF.3SG/ leave.IMPERF.3PL a student but  
 doi profesori.  
 but two professors.  
 “A student wasn’t leaving but two professors were.”

- (53) Plecau / \*pleca nu un student  
 Leave.IMPERF.3SG / leave.IMPERF.3PL not a student  
 ci doi profesori  
 but two professors  
 “A student wasn’t leaving but two professors were.”
- (54) Nu a mâncat / \*au mâncat mâncarea din frigider  
 Not have.3SG / have.3PL eaten food.DEF from fridge  
 chiriaşul ci cei doi musafiri.  
 tenant.DEF but DEF two guests  
 “The tenant didn’t eat the food in the fridge but the two guests did.”
- (55) Au mâncat / \*a mâncat mâncarea din frigider nu chiriaşul  
 Have.3PL / has.3SG eaten food.DEF from fridge not tenant.DEF  
 ci cei doi musafiri.  
 but DEF two guests.  
 “The tenant didn’t eat the food in the fridge but the two guests did.”

(52) and (54) are anchored correctives with sentential negation, which show a first conjunct agreement effect; (53) and (55) are basic correctives, where no first conjunct agreement effect is present and the agreement pattern is the same as in standard ‘and’-coordination (56):

- (56) Au mâncat / \*a mâncat mâncarea din frigider chiriaşul  
 Have.3PL eaten / have.3SG food.DEF from fridge tenant.DEF  
 şi cei doi musafiri.  
 and DEF two guests  
 “The tenant and the two guests have eaten the food in the fridge.”

Similar effects obtain with depictive secondary predication (57-58) and passives (59-60) (see Franco 2016 for Italian data):

- (57) Nu mănânc ardeiul ci roşia crudă /  
 Not eat.1.SG chilly.DEF.MASC but tomato.DEF.FEM raw.FEM/  
 \*crud / \*cruzi.  
 raw.MASC/ raw.PL  
 “I don’t eat the chilly but the tomato raw.”

- (58) Mănânc nu ardeicul                      ci roșia                      crudă/  
 Eat.1.SG not chilly.DEF.MASC but tomato.DEF.FEM                      raw.FEM/  
 \*cruzi/\*crud.  
 raw.MASC.PL / raw.MASC.SG  
 “I don’t eat the chilly but the tomato raw.”
- (59) Nu sunt certate / \*certați                      fetele  
 Not are scolded.FEM.PL/ scolded.MASC.PL girls.DEF.FEM  
 ci băieții  
 but boys.DEF.MASC.  
 “The girls are not scolded but the boys are.”
- (60) Sunt certați /                      ?? certate                      nu fetele,                      ci băieții.  
 Are scolded.MASC.PL/ scolded.FEM.PL not girls.DEF but boys.DEF  
 “The girls are not scolded but the boys are.”

To briefly conclude the section, agreement data in Romanian point to an analysis of anchored corrective sentences as involving clausal coordination followed by ellipsis, and of basic corrective sentences as involving subclausal coordination.

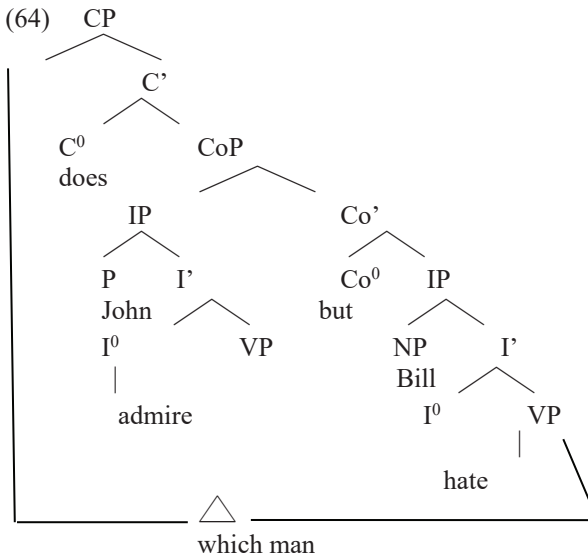
## 6. Steps towards a unifying analysis; further research

Adopting any of the available theories on multidominance (de Vries 2005, 2009, 2013, Citko 2012, van Riemsdijk 2000, 2001, 2006) may turn out to dispense with the need for positing different syntactic structures for anchored and basic corrective sentences.

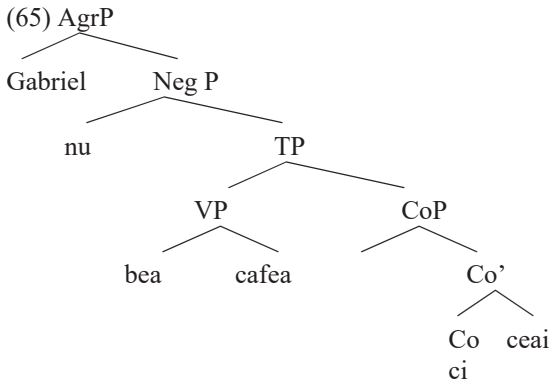
For example, de Vries (2009) derives multidominance structures via the operation of external remerge defined as a natural output of merge (see also van Riemsdijk 2000, on grafts as the natural output of merge). External remerge derives structures where a constituent is moved to an independent structure, resulting in the sharing of that constituent between structures.

Examples of external remerge are RNR (61), wh-amalgams/cleft-amalgams (62), Across-the-Board Movement (63):

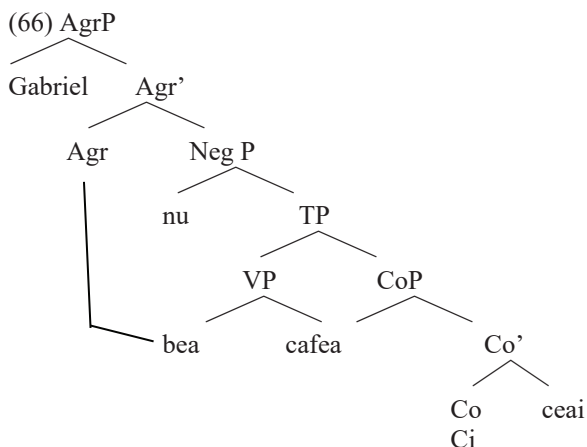
- (61) John admires \_\_\_\_, but Bill hates Trump.  
 (62) Jack gave [you will never guess which girl] a flower  
 John gave [I think it was his girlfriend] a flower  
 (63) Which man does John \_\_ admire but Bill hate?



An attempt at applying a multidominant approach to the structure of anchored correctives would look like (65):



In turn, a multidominant structure for basic correctives would look like (66), where V ‘bea’ raises to Agr:



## 7. Conclusions

The paper has shown that Romanian has two ‘but’s: ‘*ci*’ for the corrective reading and ‘*dar*’ for the counterexpectational reading. It has also shown that the syntax of corrective ‘*ci*’ sentences in the anchored form suggests an analysis in terms of clausal coordination followed by ellipsis (apud Vicente 2010). In turn, the syntax of corrective ‘*ci*’ sentences in the basic form suggests an analysis in terms of subclausal coordination (apud Toosarvandani 2013). Further research will show if multidominance can shed light on the syntax of corrective sentences.

## References

- Abraham, Werner 1979. “But”. *Studia Linguistica* 33 (2): 89-119.
- Anscombe, Jean-Claude and Oswald Ducrot. 1977. “Deux mais en français?”. *Lingua* 43: 23-40.
- Bianchi, Valentina, and Roberto Zamparelli. 2004. “Edge coordinations: focus and conjunction reduction”. In *Peripheries: Syntactic Edges and Their Effects*, edited by David Adger, Cécile de Cat, and George Tsoulas, 313-328. Dordrecht: Kluwer.
- Bosque, Ignacio. 1980. *Sobre la negación*. Madrid: Cátedra.
- Citko, Barbara. 2012. “Multidominance”. In *The Oxford Handbook of Linguistic Theory*, edited by Cedric Boeckx, 119-143. Oxford: Oxford University Press.

- Dascal, Marcelo, and Tamar Katriel. 1977. "Between semantics and pragmatics: the two types of 'but' - Hebrew 'aval' and 'ela'". *Theoretical Linguistics* 4 (3): 143-172.
- de Vries, Mark. 2005. "Coordination and syntactic hierarchy". *Studia Linguistica* 59 (1): 83-105.
- de Vries, Mark. 2009. "On multidominance and linearization". *Biolinguistics* 3 (4): 344-403.
- de Vries, Mark. 2013. "Multidominance and locality". *Lingua* 134: 149-169
- Franco, Ludovico. 2016. "Adversative "corrective" coordination: further evidence for combining sub-clausal constituents". *Revue Roumaine de Linguistique* LXI (2): 125-142.
- Horn, Laurence R. 1989. *A natural history of negation*. Stanford: CSLI Publications.
- Langacker, Ronald W. 1969. "On pronominalization and the chain of command". In *Modern Studies in English*, edited by David A. Reibel, and Sanford A. Schane, 160-186. Englewood Cliffs: Prentice Hall.
- McCawley, James D. 1991. "Contrastive negation and metalinguistic negation". In *CLS 27: the Parasession on Negation*, edited by Lise M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez, 189-206. Chicago: Chicago Linguistic Society.
- Riemsdijk, Henk van. 2000. "Free Relatives Inside Out. Transparent Free Relatives as Grafts". In *PASE Papers in Language Studies. Proceedings of the 8th Annual Conference of the Polish Association for the Study of English*, edited by Bożena Rozwadowska, 223-233. Wrocław: University of Wrocław.
- Riemsdijk, Henk van. 2001. "A Far from Simple Matter. Syntactic Reflexes of Semantics-Pragmatics Misalignments". In *Semantics, Pragmatics and Discourse. Perspectives and Connections. A Festschrift for Ferencz Kiefer*, edited by Istvan Kenesei, and Robert M. Harnish, 21-41. Amsterdam: John Benjamins.
- Riemsdijk, Henk van. 2006. "Grafts Follow from Merge". In *Phases of Interpretation*, edited by Mara Frascarelli, 17-44. Berlin: Mouton de Gruyter.
- Rizzi, Luigi. 1997. "The fine structure of the left periphery". In *Elements of grammar*, edited by Liliane Haegeman, 281-337. Dordrecht: Kluwer.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. PhD diss., MIT.
- Ross, John Robert. 1969. "Guess who?". In *Proceedings from the Annual Meeting of the Chicago Linguistic Society 5*, edited by Robert I.

- Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252-286. Chicago: University of Chicago, Chicago Linguistic Society.
- Steindl, Ulrike. 2013. "Corrective coordination with *but* across languages". Talk presented at CGSW 28, October 4th 2013.
- Toosarvandani, Maziar. 2013. "Corrective *but* coordinates clauses not always *but* sometimes". *Natural Language and Linguistic Theory*, 31 (3): 827-863.
- Toosarvandani, Maziar. 2014. "Contrast and the structure of discourse". *Semantics and Pragmatics*, 7 (4): 1-57.
- Van der Wouden, Ton. 1997. *Negative Contexts*. PhD diss., University of Groningen.
- Vicente, Luis. 2010. "On the syntax of adversative coordination". *Natural Language and Linguistic Theory* 28 (2): 381-415.
- Zafiu, Rodica. 2005. "Conjuncțiile adversative din limba română: tipologie și niveluri de incidență". In *Limba română – structură și funcționare*, edited by Gabriela Pană Dindelegan, 243-258. Bucharest: Editura Universității din București.

## Note

---

<sup>1</sup> Toosarvandani (2013) argues that the flat structure of the conjunction phrase in (32) is preferred for reasons of simplicity

# CHAPTER THIRTEEN

## SUBORDINATION AND FRAGMENTS: THE CASE OF ROMANIAN

GABRIELA BÎLBÎIE

In this paper we investigate three different constructions (namely polar verbless clauses, stripping and gapping) in order to show that, despite the syntactic and semantic differences between them, they display a similar behaviour with respect to embedding in Romanian, i.e., they all allow embedding in some very specific contexts. We discuss the semantic factor of factivity and we show that the classical distinction between non-factives and factives is too reductive and cannot fully account for all the empirical facts we observe. Factives are not all alike: in particular, we observe that semi-factives come closer to non-factive predicates rather than to other factive predicates, when it comes to fragments embedding. The similar behaviour of semi-factive and non-factive predicates in embedding fragments challenges the ‘size-of-complement’ approach (which syntactically explains the semantic asymmetry between non-factives and factives), and calls on alternative explanations in semantic and pragmatic terms.

Keywords: polar verbless clauses, stripping, gapping, embedding, factivity, Romanian

### 1. Introduction

It is usually assumed that some elliptical constructions (such as gapping, cf. Hankamer 1979, Johnson 2009, 2014, 2018) are root phenomena, being excluded in embedded contexts. In this paper, we investigate three different constructions, namely polar verbless clauses (henceforth PVCs), stripping and gapping, in Romanian embedded contexts. In English, all these constructions are considered to be ungrammatical under embedding (Johnson



2009, 2018, Weir 2014, Wurmbrand 2017), as illustrated in (1) for PVCs, in (2) for stripping, and in (3) for gapping.

- (1) a. A: Is John coming to the party? B: \***I think that** {yes/no}.
- b. \*John is coming to the party, but **I think that** Mary no.
- (2) a. A: Who left? B: \***I think that** Bill.
- b. \*John drinks scotch, and **I think that** Bill too.
- (3) \*John drinks scotch, and **I think that** Bill bourbon.

In other languages, this is subject to variation: for example, French (Authier 2013, Abeillé *et al.* 2014, Pasquereau 2018) allows embedding in PVCs (4), in polar stripping (5b), but not in non-polar stripping (5a) or gapping (6). On the other hand, Romanian allows embedding in all these constructions (7-9), like Spanish (10-12) (de Cuba & MacDonald 2013, García-Marchena 2018, Bîlbîie & de la Fuente 2019), but in very specific contexts.

- (4) a. A: Est-ce que Jean viendra à la fête? B: **Je pense que** {oui/non}.<sup>1</sup>  
 ‘A: Will Jean come to the party? B: I think that {yes/no}.’
- b. Jean viendra à la fête, mais **je pense que** Marie non.  
 ‘Jean will come to the party, but I think that Marie no.’
- (5) a. A: Qui est parti? B: \***Je pense que** Jean.  
 ‘A: Who left? B: I think that Jean.’
- b. Paul a commandé une bière et **je pense que** Jean aussi.  
 ‘Paul ordered a beer and I think that Jean too.’
- (6) \*Paul a commandé une bière et **je pense que** Jean un whisky.  
 ‘Paul ordered a beer and I think that Jean a whisky.’
- (7) a. A: Va veni Ion la petrecere? B: **Cred că** {da/nu}.
- ‘A: Will Ion come to the party? B: I think that {yes/no}.’
- b. Ion va veni la petrecere, dar **cred că** Maria nu.  
 ‘Ion will come to the party, but I think that Maria no.’
- (8) a. A: Cine a plecat? B: **Cred că** Ion.  
 ‘A: Who left? B: I think that Ion.’
- b. Paul a comandat o bere și **cred că** și Ion.  
 ‘Paul ordered a beer and I think that also Ion.’

- (9) Paul a comandat o bere și **cred că** Ion un whisky.  
‘Paul ordered a beer and I think that Ion a whisky.’
- (10) a. A: ¿Vendrá Juan a la fiesta? B: **Creo que** {sí/no}.  
‘A: Will Juan come to the party? B: I think that {yes/no}.’  
b. Juan vendrá a la fiesta pero **creo que** María no.  
‘Juan will come to the party but I think that María no.’
- (11) a. A: ¿Quién se ha ido? B: **Creo que** Juan.  
‘A: Who left? B: I think that Juan.’  
b. Pablo pidió una cerveza y **creo que** Juan también.  
‘Pablo ordered a beer and I think that Juan too.’
- (12) Pablo pidió una cerveza y **creo que** Juan un whisky.  
‘Pablo ordered a beer and I think that Juan a whisky.’

In this paper, we focus on Romanian data. The main goal of the paper is to show that, despite the syntactic and semantic differences between these three constructions, they display a similar behaviour with respect to embedding.

The paper is organized as follows: In Section 2, we briefly present the three constructions in their regular uses (without embedding), in order to show that the ellipsis process is not necessarily involved in all of them; in particular, PVCs are not elliptical at all. In Section 3, we show that, despite the differences we observe between these three constructions, there is a general semantic constraint at work in all of them. In Section 4, we propose an explanation for the semantic constraint observed in Section 3.

## 2. Three different constructions

We begin this section by defining the three constructions under investigation: PVCs, stripping and gapping.

**PVCs** (or pseudostripping, cf. Depiante 2000) make use of polar propositional adverbs such as *da* ‘yes’ and *nu* ‘no’ in Romanian. They occur, in particular, as short answers (cf. Sailor 2012, ‘polar response particles’) in dialogic contexts (13), but they can also occur in monologic contexts, in so-called contrastive coordinations (14). Polarity particles *da* and *nu* can be bare (13a, 14a), or preceded by a topic XP (13b, 14b-c).

- (13) a. A: Va veni Ion la petrecere? B: {**Nu/Da**}.  
‘A: Will Ion come to the party? B: {No/yes}.’

- b. A: *Îți vor veni copiii de Paște?* B: **Ion da**, dar **Maria nu**.  
‘A: Will your children come for Easter? B: Ion yes, but Maria no.’
- (14) a. *Astăzi se decide dacă se închid (sau nu) școlile (sau nu).*  
‘Today one decides whether schools close (or not).’
- b. *Ion va veni la petrecere, dar Maria nu.*  
‘Ion will come to the party, but Maria no.’
- c. *Ion nu va veni la petrecere, dar Maria da.*  
‘Ion will not come to the party, but Maria yes.’

**Stripping** (or Bare Argument Ellipsis, cf. Ross 1969, Hankamer & Sag 1976) refers to any elliptical clause displaying a single remnant which frequently – but not necessarily – contains a focus-sensitive adverb, such as the polar additive particle *și*<sup>2</sup> ‘too’ (16a), the polar restrictive particle *nici* ‘neither’ (15b, 16b) or the constituent negation *nu*<sup>3</sup> ‘not’ (17). Stripping elements can appear as answers in a dialogue (also called fragment answers) as in (15), or they can be coordinated (16-17).

- (15) a. A: *Cine a venit?* B: **Ion**.  
‘A: Who came? B: Ion.’
- b. A: *Nu mi-am făcut tema.* B: **Nici eu**.  
‘A: I did not do my homework. B: Neither me.’
- (16) a. *Vine Ion la petrecere, dar și Maria.*  
‘Ion is coming to the party, and also Maria.’
- b. *Ion nu vine la petrecere, și nici Maria.*  
‘Ion is not coming to the party, and neither Maria.’
- (17) a. *Profesorul le-a acordat altora a doua șansă, dar nu mie.*  
‘The teacher gave others a second chance, but not to me.’
- b. *Femeile au un cuvânt de spus, și nu bărbații.*  
‘The women have a word to say, and not the men.’

The adverbs which may accompany the stripped phrase are called associative adverbs, or focus-sensitive particles: they associate with a focused constituent, which in stripping cases is the remnant. Syntactically, they behave as phrase modifiers. Semantically, they behave as operators which quantify over alternatives varying in the focused position. This allows us to distinguish between the proform negation *nu*<sub>1</sub> in PVCs (18a), and the constituent negation *nu*<sub>2</sub> in stripping (18b). Their position in the

clause is linked to a different informational status and interpretation: the proform negation in (18a) follows a contrastive topic, whereas the constituent negation in (18b) precedes a focused element. The sequences containing these negative adverbs do not give rise to the same interpretation either, as shown by the different paraphrases in (18): the dative phrase *mie* ‘to me’ in (18a) behaves as a topic (cf. the ‘aboutness’ feature highlighted by the conventionalized syntactic structure ‘as for X’), whereas the dative *mie* ‘to me’ in (18b) is rather a focus (cf. its use in a cleft structure of type ‘it is X that’). Moreover, the proform negation has propositional scope in (18a), whereas the constituent negation does not (i.e., the target of the negation only corresponds to the content of its associate element).

- (18) a. Profesorul le-a acordat altora a doua șansă, dar **mie nu**. (= ‘As for me, the teacher didn’t give me a second chance.’)  
 ‘The teacher gave others a second chance, but to me no.’
- b. Profesorul le-a acordat altora a doua șansă, dar **nu mie**. (= ‘It is not to me that the teacher gave a second chance.’)  
 ‘The teacher gave others a second chance, but not to me.’

**Gapping** (Ross 1967) can be generally defined as any elliptical clause containing at least two remnants (one of them being generally – but not necessarily – the subject) and lacking at least the main verb. Unlike the stripped clause which contains only a remnant (modified in some cases by an associative adverb), the gapped clause contains two or more remnants, which are paired with some correlates in the non-elliptical clause (i.e. the source). Most cases of gapping occur in coordination contexts (19a), but it may occur in dialogue too (19b-c).<sup>4</sup> In some cases, the gapped clause may contain also an adverbial modifier, like in stripping cases, such as the associative adverb *nici* in (20).

- (19) a. Paul a comandat o bere {și/iar} **Ion un whisky**.  
 ‘Paul ordered a beer and Ion a whisky.’
- b. A: Cine ce a comandat? B: **Paul o bere, (și/iar) Ion un whisky**.  
 ‘A: Who ordered what? B: Paul a beer, (and) Ion a whisky.’
- c. A: Eu vreau să merg la mare. B: Iar **eu la munte**.  
 ‘A: I want to go to the sea. B: And I to the mountain.’
- (20) Ion nu merge la film și **nici Maria la teatru**.  
 ‘Ion does not go to the cinema and neither Maria to the theater.’

One may say that stripping is a subtype of gapping (cf. Hankamer & Sag 1976 and the subsequent literature). If this is the case, one expects that it has the same properties as gapping. However, the different number of remnants is correlated with different informational status of these two constructions (Bilbiie 2017): the stripped phrase is a contrastive focus (CF), whereas the gapped clause has a more complex information structure: in most cases, a contrastive topic (CT) followed by a contrastive focus (CF), as shown by the contrast in (21). Moreover, this difference in information structure contribution is also supported by the conjunctions use: the topic-marking conjunction *iar* (which always introduces at least two phrases, the first one being necessarily a contrastive topic) is incompatible with stripping (21b), but is by far the most used conjunction in gapping (21a).<sup>5</sup>

- (21) a. [Ion]<sub>CT</sub> joacă [volei]<sub>CF</sub>, iar [Maria]<sub>CT</sub> [tenis]<sub>CF</sub>.  
 ‘Ion plays volleyball, and Maria tennis.’  
 b. Ion joacă [volei]<sub>CF</sub>, {dar/\*iar} [și tenis]<sub>CF</sub>.  
 ‘Ion plays volleyball, but also tennis.’

The common aspect bringing these three constructions together is this idea that there is interpretation beyond what is said/written, something is literally missing, or is semantically much less contentful than what is actually understood, and that what is understood is understood because of the presence of an antecedent in the context. Furthermore, a uniform analysis is usually assumed for describing their syntactic behaviour, i.e., all three involve some elliptical process, in particular clausal ellipsis.

However, we consider that there are two mechanisms at work: proform-analysis vs. ellipsis-analysis. This recalls the very influential paper of Hankamer & Sag (1976) which distinguishes between two main classes of anaphoric devices: deep vs. surface anaphora. Deep anaphora do not result from an ellipsis process, they are rather model-interpretive anaphora (e.g. *do it, do this, do that*, personal pronouns). On the other hand, surface anaphora result from an ellipsis mechanism (such as stripping or gapping). According to Hankamer and Sag’s dichotomy, surface anaphora must be linked to a linguistic (syntactically present) antecedent (i.e. they only have endophoric uses), whereas deep anaphora do not necessarily require a linguistic antecedent (i.e. they allow both endophoric and exophoric uses). If one comes back to our three constructions, we observe that PVCs may have pragmatic antecedents, i.e. they may occur in situations in which the antecedent is presented in the context, but not introduced explicitly in a linguistic expression (‘pragmatic control’, as in (22)). Therefore, unlike elliptical clauses (stripping or gapping), PVCs can have an exophoric use

(with no linguistic antecedent), which is interpreted either through ostension or via inference from the non-linguistic context.

- (22) [*Context: Speaker A raises the teapot and wants to serve B.*]
- a. B<sub>1</sub>: [Nu], mulțumesc.  
'No, thanks.'
  - b. B<sub>2</sub>: [Acum nu], poate mai târziu.  
'Not now, maybe later.'

Moreover, Tanenhaus & Carlson (1990) show clear evidence for an interaction between syntactic parallelism and type of anaphor: surface anaphora are sensitive to the form of their antecedents in a way that deep anaphora are not. The adverbial proforms in PVCs are not sensitive to the form of their antecedents: they may substitute not only finite clauses or verbal phrases, but also other kinds of phrases, provided that they have a predicative use; the form that their antecedent may have is therefore highly underspecified. In (23), the adverbial proform *nu* may substitute a nominal, adjectival or prepositional phrase which, in some cases, cannot have a lexical negated form (*\*nestudent* / *\*nu student*; *\*netânăr* / *\*nu tânăr*), challenging the syntactic reconstruction approach. In such cases, strict reconstruction of an 'antecedent' into an ellipsis site is not conducive to acceptable results.

- (23) a. [Student [sau nu]], [tânăr [sau nu]], oricine e binevenit în comunitatea Euforia.  
'Student or not, young or not, anyone is welcome in the Euforia community.'
- b. [Cu profesor [sau nu]], ei se consacrau cu mare plăcere studiului.  
'With a professor or not, they devoted themselves with great pleasure to the study.'

Therefore, we analyze PVCs as instances of deep anaphora, whereas stripping and gapping are rather instances of surface anaphora (see also Johnson 2018). Moreover, we adopt a proform-analysis (Krifka 2013) of polarity particles *da* 'yes' and *nu* 'no' in PVCs, rather than an elliptical one (Kramer & Rawlins 2011, Holmberg 2015). They are not the remnant of ellipsis (as assumed by Kramer & Rawlins 2011 for English, de Cuba & MacDonald 2013 for Spanish, Authier 2013 and Pasquereau 2018 for French), but rather proforms. These polarity adverbs behave as propositional anaphors: they partly receive their interpretation through a contextually

given antecedent (like an anaphorical pronoun). PVCs as a whole are therefore clauses with a non-verbal predicative head. In the case at stake, the predicative head is an adverbial phrase containing a propositional adverb such as *da* or *nu* in Romanian. PVCs can contain only the predicative head (the adverbial phrase itself), or two phrases, the predicative adverbial phrase being preceded by a topic phrase (hanging topic, cf. Krifka 2013).

On the other hand, stripping and gapping constructions are elliptical clauses: they have a propositional content, but they lack the predicative verbal head. While this goes beyond the scope of this paper, we note that there are two main routes to account for the internal syntax of these elliptical clauses:<sup>6</sup> the first appealing to a syntactic reconstruction mechanism, involving some deletion process (à la Merchant 2004), and the second appealing to a semantic reconstruction mechanism and a dedicated meaning-form rule that links a headless structure (i.e. a syntactic fragment) to a clausal meaning (Ginzburg & Sag 2000, Abeillé *et al.* 2014, Bilbîie 2017).

### 3. Constraints on embedding

In this section, we look at the behaviour of PVCs, stripping and gapping in embedded contexts. By and large, this empirical domain has not been looked at in detail. The first aspect to note is the fact that, despite the differences we have observed between PVCs, stripping and gapping in Section 2, they seem to have the same behaviour under embedding in Romanian, cf. examples (7-9) above: namely, they can occur in embedded contexts.

Based on English data, such as (1-3) above, scholars have assumed that the three constructions are unable to occur in embedded contexts (see, e.g., Sailor 2012 for PVCs, Wurmbrand 2017 for stripping and Johnson 2009 for gapping). These observations sometimes led to strong generalizations, such as Hankamer (1979)'s 'Downward Bounding' constraint or Johnson (2014)'s 'No Embedding Constraint', formalized in (24), which state that stripping and gapping are root phenomena. However, a closer look at the data from a cross-linguistic perspective shows a more nuanced picture.

- (24) Downward Bounding (Johnson 2018)  
 “Let  $\alpha$  be some member of the verbal sequence of the right conjunct, and  $\beta$  be the set of elements in the sequence that c-command  $\alpha$ . If Gapping or Stripping includes  $\alpha$  then it must include  $\beta$ .”

First, one observes that, cross-linguistically, embedding is sensitive to the presence or absence of the complementizer. In English, embedded stripping (25a) or gapping (25b) may occur in certain contexts if the complementizer is absent (i.e., *that*-less clauses). See, in this regard, the Embedded Stripping Generalization proposed by Wurmbrand (2017: 345) for English: “Stripping of embedded clauses is only possible when the embedded clause lacks a CP.” She extends it to gapping as well, by noticing that omission of the complementizer “should also show an ameliorating effect in gapping” (Wurmbrand 2017: 361).

- (25) a. Jane loves to study rocks, and John says (\*that) geography too. (Wurmbrand 2017: 344)  
 b. Some will eat mussels and she claims (\*that) others shrimp. (Wurmbrand 2017: 361)

For languages licensing embedded fragments in the absence of complementizer (e.g., English-type languages: English, Dutch, German, cf. Vicente 2013), one could argue that there is no true embedding in these contexts (cf. Temmerman 2013 for embedded fragment answers): the putative embedded fragment is rather a matrix fragment, the embedding sequence to its left being rather a parenthetical (accordingly, (26a) will be parallel to (26b)). Whereas this analysis is questionable even for English (see Weir 2014’s discussion), it cannot be applied to languages which require an obligatory complementizer, such as Romanian (but also Spanish, Polish, Czech, cf. Vicente 2013). In these languages, there is true embedding in these contexts, since the presence of complementizer is not possible in constructions with a parenthetical (compare the parenthetical use of the verb *cred* ‘I believe’ in (27b) in Romanian and the true embedding use of the same verb in (27a)).

- (26) Who left?  
 a. I think (\*that) John. (Weir 2014: 212)  
 b. John I think. (Weir 2014: 219)
- (27) Cine a mâncat prăjitura?  
 ‘Who ate the cake?’  
 a. Cred \*(că) Ion; el adoră dulciurile.  
 ‘I think that Ion; he loves sweets.’  
 b. Ion, cred \*(că); el adoră dulciurile.  
 ‘Ion, I think; he loves sweets.’



Second, if embedding is involved (be it with or without a complementizer), one observes that it is constrained by the semantic type of the embedding predicate. This was first noticed by de Cuba & MacDonald (2013) and Fernández-Sánchez (2017) for Spanish, who observe that embedding is possible with non-factive verbs, but impossible with factive verbs. The contrast is illustrated in (28) for PVCs, in (29) for stripping and in (30) for gapping. The same contrast between two classes of predicates is observed in English by Sailor (2012) for PVCs, and Weir (2014) for stripping and gapping.<sup>7</sup> Therefore, all previous studies taking into account the factivity factor distinguish between non-factive predicates, that allow embedded PVCs and fragments (e.g., *suppose*, *imagine*, *suspect*, *think*, *say*, etc.), and factive predicates, that do not allow embedded PVCs and fragments (e.g. *hate*, *love*, *know*, *find out*, *regret*, etc.).

- (28) ¿Llegaron a tiempo? (de Cuba & MacDonald 2013: 312)  
 ‘Did they arrive on time?’  
 a. {Creo / me parece} que {sí/no}.  
 ‘{I think / it seems} that {yes/no}.’  
 b. \*{Lamento / me desagrada} que {sí/no}.  
 ‘{I regret / it displeases me} that {yes/no}.’
- (29) ¿Quién robó las joyas? (de Cuba & MacDonald 2013: 321)  
 ‘Who stole the jewels?’  
 a. {Creo / supongo / me imagino / pienso} que tu hijo.  
 ‘{I believe / I suppose / I imagine / I think} that your son.’  
 b. #{Lamento / sé / me sorprende / me desagrada} que tu hijo.  
 ‘{I regret / I know / it surprises me / it displeases me} that your son.’
- (30) a. Alfonso robó las esmeraldas y {creo / imagino / supongo / ...}  
 que Mugsy las perlas. (Fernández-Sánchez 2017: 10)  
 ‘Alfonso stole the emeralds and {I think / I imagine / I  
 suppose / ...} that Mugsy the pearls.’  
 b. \*Alfonso robó las esmeraldas y {lamento / me encanta / odio  
 / ...} que Mugsy las perlas.  
 ‘Alfonso stole the emeralds and {I regret / I love / I hate / ...}  
 that Mugsy the pearls.’

However, a closer look at data shows that this binary distinction between non-factive and factive predicates is too reductive and cannot fully account for all the empirical facts. All previous studies on embedded fragments do

not take into account a more fine-grained distinction between types of predicates (as proposed by Karttunen 1971, Kiparsky & Kiparsky 1971, Hooper 1975). In particular, they miss the heterogeneous behaviour of factive verbs; some factive verbs, namely knowledge predicates (such as *know*, *find out*, *see*, *notice*, *discover*), are supposed to be unacceptable: e.g., the verb *sé* ‘I know’ in (29b) above, which is assumed to behave as other factive predicates. Crucially, this subclass of factive verbs is nevertheless attested in some naturalistic (web and corpus) data.

Therefore, we propose to distinguish between three types of embedding predicates (Karttunen 1971): non-factive predicates (such as epistemic and communication verbs, e.g., *suppose*, *imagine*, *suspect*, *think*, *believe*, etc.), semi-factives (such as knowledge predicates, e.g., *know*, *find out*, *see*, *notice*, *discover*) and true factives (such as emotion predicates, e.g., *regret*, *like*, *resent*, *be surprised*, etc.).

In Romanian (as well as in Spanish, cf. Bîlbîie & de la Fuente 2019), PVCs and fragments can be embedded not only under non-factive predicates, but also under semi-factive ones. However, there are no attested examples with embedding under true factive predicates.

Romanian PVCs (containing the propositional adverbs *da* ‘yes’ or *nu* ‘no’)<sup>8</sup> seem to be the most frequent embedded structures, compared to embedded stripping or embedded gapping. They can easily be embedded under non-factive verbs (31), such as *se pare* ‘it seems’ (31a), *bănuiesc* ‘I guess’ (31b), *presupun* ‘I suppose’ (31c-d). Crucially, they may also be embedded under semi-factive verbs (32), such as *știu* ‘I know’ (32a), *văd* ‘I see’ (32b), *observ* ‘I observe’ (32c), *am constatat* ‘I noticed’ (32d). These embedded polarity particles are propositional anaphors<sup>9</sup> (as discussed in Section 2), as shown by the presence of the same complementizer as in finite clauses (*că* ‘that’). These empirical data, in particular semi-factive contexts, strongly challenge the semantic generalization of previous studies on embedded polarity particles (de Cuba & MacDonald 2013 for Spanish, Authier 2013 and Pasquereau 2018 for French), according to which embedded PVCs are possible only under non-factive predicates.

- (31) a. Există super eroi printre noi? **Se pare că** da.  
 ‘Are there any super heroes among us? It seems that yes.’  
 b. Ați auzit de poetul bistrițean Nicolae Avram? **Bănuiesc că** nu.  
 ‘Did you hear about the Bistrița poet Nicolae Avram? I guess that no.’

- c. Ascultați multă muzică? Soțul dvs. e muzician, deci **presupun că da**.  
'Do you listen to a lot of music? Your husband is a musician, so I suppose that yes.'
- d. Nu știu sigur dacă în acest caz vei plăti din nou taxa, dar **presupun că nu**.  
'I'm not sure if in this case you will pay the tax again, but I suppose that no.'
- (32) a. Folosește Alex Velea steroizi? **Eu știu că nu**.  
'Does Alex Velea use steroids? I know that no.'
- b. Poate fi ucisă dragostea? **Eu văd că da**. Ții la dragoste mai mult decât la tine însuși? **Eu văd că nu**.  
'Can love be killed? I see that yes. Do you like love more than yourself? I see that no.'
- c. Vreți să promovăm prin metode nemeritate? **Observ că da**.  
'Do you want us to promote through unmerited methods? I observe that yes.'
- d. Am verificat dacă supapa de la recirculare funcționează și **am constatat că da**.  
'I checked if the recirculation valve is working and I noticed that yes.'

Moreover, Romanian allows embedding of any stripping construction, be it a fragment answer or a polar stripped fragment in a coordination context (unlike French, cf. the contrast in (5) and (8) above). Just as in the case of PVCs, embedded stripping is possible not only under non-factive predicates (33), but also under semi-factive verbs (34). Weir (2014)'s generalization on fragments embedding is based on a syntactic licensing of ellipsis: only 'bridge' verbs (i.e., non-factive predicates) can embed fragments. However, the fact that semi-factive verbs can embed stripping in Romanian challenges Weir's generalization, as semi-factive verbs are generally analyzed as 'non-bridge' verbs.

- (33) a. Ne-am plăcut din prima clipă când ne-am privit, a fost dragoste la prima vedere! Nu mi s-a întâmplat niciodată, și **cred că nici lui!**  
'We loved each other from the first moment when we met, it was love at first sight! It never happened to me, and I think that to him neither!'

- b. Sunt bine, **sper că** și tu.  
'I am fine, I hope that you too.'
- c. Cine a citit toate cărțile? **Bănuiesc că** nimeni.  
'Who read all the books? I guess that nobody.'
- (34) a. Eu nu cred în chestii din astea fancy și **știu că** nici tu.  
'I don't believe in these fancy things and I know that you neither.'
- b. Imi plac mult florile, și **văd că** și ție.  
'I like much the flowers, and I see that you too.'
- c. Nu știu cât de bine s-ar descurca, dar **știu că** nu foarte bine.  
'I don't know how well he would do, but I know that not very well.'

Although embedded gapping is less frequent than embedded PVCs and stripping in Romanian, it follows the same pattern: embedded gapping is acceptable under non-factive (35) and semi-factive (36) verbs. The fact that all our naturalistic data in (35) and (36) involve some symmetric/reciprocal relation between the source and the gapped clause is not surprising at all, since gapping is generally assumed to involve two contrastive pairs (for more details, see Bilbîie 2017).

- (35) a. Am să-l țin minte toată viața pe Gigi, și **cred că** și el pe mine, mărturisește Florin.  
'I will remember all my life Gigi, and I think that he me too, confesses Florin.'
- b. Amintiri frumoase pe care mi le-au lăsat ei mie, și **sper că** și eu lor.  
'Beautiful memories that they left me, and I hope that I them too.'
- c. Nu eu îl urăsc pe el, ci **cred că** el pe mine.  
'I don't hate him, but I think that he me.'
- (36) a. Ion o iubește pe Ana și **văd că** și ea pe el.  
'Ion loves Ana and I see that she him too.'
- b. Nu îmi imaginez viața fără părinții mei; îi iubesc enorm de mult și **știu că** și ei pe mine.  
'I can't imagine life without my parents; I love them a lot and I know that they me too.'

Recent experimental work on embedded gapping from a cross-linguistic perspective (Bilbâie & de la Fuente 2019, Bilbâie *et al.* 2019) show, based on acceptability judgment tasks for Spanish, Romanian, French, and English, that there is cross-linguistic variation with respect to embedded gapping; embedded gapping is acceptable in Romanian and Spanish, and less so in French and English, but it obeys more general semantic constraints: non-factive verbs embed more easily than factive ones, and within factive predicates, semi-factive verbs embed better than true factive ones. Crucially, despite the cross-linguistic variation, factivity is a significant factor in all languages: embedded clauses under a factive verb are less acceptable than under a non-factive verb, and embedded clauses under a true factive verb are less acceptable than under a semi-factive verb.

Based on Romanian data, we conclude that embedding of a PVC, a stripped or a gapped clause is possible, which shows the limitations of previous generalizations such as the ‘No Embedding Constraint’ (Johnson 2014): in particular, the ‘No Embedding Constraint’ is not as strict and universal as traditionally assumed. Moreover, there is true embedding as shown by the obligatory presence of the complementizer in Romanian. Finally, there is sensitivity to the semantic class of embedding predicate: some predicates easily allow embedding, whereas others do not. We have shown that the classical distinction between non-factives and factives is too reductive and cannot fully account for all the empirical facts we observe. Factives are not all alike: in particular, we observe that semi-factives come closer to non-factive predicates rather than to other factive predicates, when it comes to fragments embedding.

#### 4. Towards an explanation

The fact that not all verbs allow embedding of fragments could receive an explanation in syntactic terms. Weir (2014) and Fernández-Sánchez (2017) propose such an explanation, based on the widespread idea that non-factive and factive verbs (or ‘bridge’ vs. ‘non-bridge’ verbs) do not involve the same syntactic structure, namely factive verbs display a less complex syntactic structure than non-factives (cf. Haegeman 2006). Therefore, Weir (2014) argues that embedded fragments under non-factives (or ‘bridge’ verbs, in his terms) are ‘bigger’ than embedded fragments under factives (or ‘non-bridge’ verbs). The former contain a syntactically more complex complementizer domain (i.e. double-complementizer structure: a ‘higher’ head and a ‘lower’ one) than the latter, which display a single-complementizer structure (with no ‘high’ complementizer head). The generalization in this kind of approaches is that the verbs which support fragment embedding are

those which contain the ‘higher’ complementizer, whereas those which contain only the ‘lower’ complementizer cannot embed fragments.

However, this generalization based on syntactic explanations wrongly predicts the unacceptability of semi-factives in embedded contexts. In this kind of approaches, the only distinction which is made is between non-factives and factives. Semi-factives are assumed to behave exactly as any other factive predicate. The empirical data we discussed in Section 3 show that semi-factives, unlike other factive predicates, can nevertheless embed any kind of fragment in Romanian. These facts are not covered by the generalization above (since semi-factive verbs lack a ‘higher’ complementizer) and thus remain unexplained (for more details on the limits of such an approach, see Bilbîie & de la Fuente 2019). This kind of approach has to be revisited in order to take into account our data with semi-factive predicates.

We propose that the effects we observe in syntax between non-factives and factives, and within factives, between semi-factives and true factives, could come from other linguistic levels, namely from the semantic and discursive properties they have.

A classical explanation for the different behaviour of non-factives and factives is given by Karttunen (1971) and Kiparsky & Kiparsky (1971). On the one hand, factive predicates presuppose the truth of their complement, assigning to it the status of an established fact; therefore, the complement of these verbs cannot be asserted. On the other hand, non-factive predicates are not accompanied by a similar presupposition, leaving room for doubt and uncertainty, their complement being asserted. Moreover, according to Hooper & Thomson (1973) and Hooper (1975), non-factive verbs (e.g., verbs of communication, verbs of thought) can function parenthetically, the embedded complement clause being in these cases the ‘main assertion’. Semi-factive verbs can have the same behaviour (at least in some environments). A similar explanation for the common behaviour of non-factive and semi-factive predicates is given by Farkas (2003): both semantic classes of predicates, as strong intensional predicates, are assertive, i.e., their complement clause is assertively added to an epistemic context. On the other hand, true factive predicates are non-assertive, i.e., their complement clause has an evaluative component. Therefore, we can conclude that fragment embedding is possible when the complement clause is asserted. A lack of assertivity in the complement clause would then lead up to unacceptability of fragment embedding.

An additional explanation for this contrast comes from discourse level. Simons (2007) proposes to replace the notion of ‘main assertion’ used by Hooper & Thompson (1973) by a more discursive notion, namely ‘main point of utterance’ (MPU). The diagnostic for MPU is the question-answer

pair. “I assume that whatever proposition communicated by the response constitutes an answer (complete or partial) to the question is the main point of the response” (Simons 2007: 1036). Interestingly, the complement clause of a true factive verb cannot be the MPU (it is the embedding factive verb that contribute to the MPU), whereas the complement of a non-factive or a semi-factive verb can constitute the MPU (the embedding verb in this case being parenthetical, i.e., it does not change the MPU).

A similar discursive explanation could be given by using the notion of ‘Question under Discussion’ (QUD, cf. Roberts 2012/1996, Ginzburg & Sag 2000, Reich 2007): each sentence of a discourse can be viewed as an answer to an implicit question, i.e., the QUD. This QUD perspective, which makes use of the question-answer congruence to analyze discourse coherence, has been proposed as a test for the discursive appropriateness of fragments and other elliptical constructions (Reich 2007, Ginzburg 2012). We adopt this perspective by claiming that a PVC, a stripped clause or a gapped clause has to be congruent with the QUD. This QUD may be explicit (as in the case of dialogues: question-answer exchanges) or implicit (as in the case of coordinations). In coordination constructions such as gapping, the first conjunct invokes a QUD that, in turn, licenses gapping in the second conjunct (Reich 2007, Johnson 2018).

The discursive constraints applying to PVCs, stripping and gapping in their non-embedded contexts apply also to embedded configurations. Our claim is that the availability of embedding with the three constructions is related to the ability of the embedded clause to constitute an answer (or a partial answer) to the QUD (see also Johnson 2018). Following Simons (2007), we can say that embedding under a non-factive verb or a semi-factive verb does not affect the discourse coherence, whereas embedding under a (true) factive verb gives rise to a discourse incongruence. This discourse congruence/incongruence is illustrated for stripping in (37) and for gapping in (38). The answer in (37a), embedded under the non-factive verb *cred* ‘I think’, is an appropriate answer to the explicit QUD in (37), whereas the answer in (37b), where there is embedding under the true factive verb *regret* ‘I regret’, answers a different QUD, namely no longer ‘who left’, but ‘what effect did it have on the speaker’. Similarly, in (38a) both conjuncts address the same QUD (i.e., ‘who ordered what’), the non-factive verb *cred* ‘I think’ having a parenthetical content, whereas in (38b) the gapped clause embedded under the true factive verb *regret* ‘I regret’ does not answer the same QUD as the source clause.

- (37) A: Cine a plecat?  
 ‘Who left?’
- a. B<sub>1</sub>: **Cred că** Ion.  
 ‘I think that Ion.’
- b. B<sub>2</sub>: \***Regret că** Ion.  
 ‘I regret that Ion.’
- (38) a. Paul a comandat o bere și **cred că** Ion un whisky.  
 ‘Paul ordered a beer and I think that Ion a whisky.’
- b. \*Paul a comandat o bere și **regret că** Ion un whisky.  
 ‘Paul ordered a beer and I regret that Ion a whisky.’

We observe that fragment embedding is reflective of pragmatic factors concerning the status of the embedded clause in the larger discourse context. This is in line with Johnson (2018)’s remark: “a more careful investigation of the discourse [...] structure of Stripping and Gapping seems likely to reduce the number of open mysteries these interesting constructions harbor. I recommend this direction to those embarking on a Gapping and Stripping career.”

## 5. Conclusions

In this paper, we investigated three different constructions (namely polar verbless clauses, stripping and gapping), which, despite the fact that they do not share the same syntactic and semantic properties, display a similar behaviour with respect to embedding in Romanian, i.e., they all allow embedding (with a complementizer) in some very specific contexts. We built on previous work on fragment embedding by taking into account the semantic factor of factivity; however, we showed that the classical distinction between non-factives and factives is too reductive and cannot fully account for all the empirical facts we observe in Romanian. All previous studies on embedded fragments do not take into account a more fine-grained distinction between types of predicates; moreover, semi-factive verbs, such as knowledge predicates, are assumed to behave as true factive verbs, that is emotion predicates. Crucially, we observed that factives are not all alike: in particular, semi-factives come closer to non-factive predicates rather than to other factive predicates, when it comes to fragments embedding. The similar behaviour of semi-factive and non-factive predicates in embedding fragments challenges the ‘size-of-complement’ approach (which explains the semantic asymmetry between non-factives and factives in terms of different syntactic structures: structures with a



‘higher’ complementizer vs. structures with a ‘lower’ complementizer), and calls on alternative explanations in semantic and pragmatic terms, namely assertivity and discourse coherence. Therefore, we conclude that the acceptability of fragment embedding in Romanian may be related, on the one hand, to the presence of assertivity in the complement clause, and, on the other hand, to the ability of the complement clause to constitute an answer to the Question Under Discussion. Both non-factive and semi-factive predicates are assertive, and do not affect the discourse coherence, whereas true factive predicates are non-assertive, and give rise to a discourse incongruence.

## References

- Abeillé, Anne, Gabriela Bîlbiie, and François Mouret. 2014. “A Romance perspective on gapping constructions.” In *Romance perspectives on Construction Grammar (Constructional Approaches to Language)*, edited by Hans Boas and Francisco González-García, 227-267. Amsterdam/Philadelphia: John Benjamins.
- Authier, J.-Marc. 2013. “Phase-Edge Features and the Syntax of Polarity Particles”. *Linguistic Inquiry* 44 (3): 345-389.
- Barbu, Ana-Maria. 2004. “The negation *nu*: lexical or affixal item?”. In *Understanding Romanian Negation. Syntactic and Semantic Approaches in a Declarative Perspective*, edited by Emil Ionescu, 68-82. Bucharest: Editura Universităţii din Bucureşti.
- Bîlbiie, Gabriela. 2017. *Grammaire des constructions elliptiques: Une étude comparative des phrases sans verbe en roumain et en français*. Berlin: Language Science Press.
- Bîlbiie, Gabriela, and Israel de la Fuente. 2019. “Can gapping be embedded? Experimental evidence from Spanish”. *Glossa: a journal of general linguistics* 4 (1) 110: 1-39.
- Bîlbiie, Gabriela, Israel de la Fuente, and Anne Abeillé. 2019. “Embedded gapping? A cross-linguistic perspective”. Poster presented at *CUNY 2019 32<sup>nd</sup> Annual CUNY Conference on Human Sentence Processing*. University of Colorado Boulder, March 29-31, 2019.
- de Cuba, Carlos, and Jonathan E. MacDonald. 2013. “On the referential status of embedded polarity answers in Spanish”. In *Selected proceedings of the 16<sup>th</sup> Hispanic Linguistics Symposium*, edited by Jennifer Cabrelli Amaro, Gillian Lord, Ana de Prada Pérez, and Jessi Elana Aaron, 312-323. Somerville, MA: Cascadilla Press.

- Depiante, Marcela. 2000. "The syntax of deep and surface anaphora: A study of null complement anaphora and stripping/bare argument ellipsis." PhD diss., University of Connecticut at Storrs.
- Farkas, Donka. 2003. "Assertion, belief and mood choice". Paper presented at *ESSLLI, Conditional and Unconditional Modality Workshop*, Vienna, August 2003.
- Fernández-Sánchez, Javier. 2017. "Embedded gapping isn't gapping". Ms., UAB/CLT.
- García-Marchena, Oscar. 2018. "Polar verbless clauses and gapping subordination in Spanish". In *Grammar and corpora 2016*, edited by Eric Fuß, Marek Konopka, Beata Trawiński and Ulrich H. Waßner, 169-182. Heidelberg: Heidelberg University Publishing.
- Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation*. Oxford: Oxford University Press.
- Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford: CSLI Publications.
- Haegeman, Liliane. 2006. "Conditionals, factives and the left periphery". *Lingua* 116: 1651-1659.
- Hankamer, Jorge. 1979. *Deletion in coordinate structures*. New York: Garland Publishing.
- Hankamer, Jorge and Ivan A. Sag. 1976. "Deep and surface anaphora". *Linguistic Inquiry* 7(3): 391-428.
- Holmberg, Anders. 2015. *The Syntax of Yes and No*. Oxford: Oxford University Press.
- Hooper, Joan. 1975. "On assertive predicates". In *Syntax and Semantics 4*, edited by John P. Kimball, 91-124. New York: Academic Press.
- Hooper, Joan, and Sandra Thompson. 1973. "On the applicability of root transformations". *Linguistic Inquiry* 4: 465-497.
- Johnson, Kyle. 2009. "Gapping is not (VP)-Ellipsis". *Linguistic Inquiry* 40 (2): 289-328.
- Johnson, Kyle. 2014. "Gapping". Ms., University of Massachusetts, Amherst.
- Johnson, Kyle. 2018. "Gapping and Stripping". In *The Oxford Handbook of Ellipsis*, edited by Jeroen van Craenenbroeck, and Tanja Temmerman, Chapter 23. Oxford: Oxford University Press.
- Karttunen, Lauri. 1971. "Some Observations on Factivity". *Papers in Linguistics* 4: 55-69.
- Kiparsky, Paul, and Carol Kiparsky. 1971. "Fact." In *Semantics. An interdisciplinary reader in philosophy, linguistics and psychology*,

- edited by Danny D. Steinberg, and Leon A. Jakobovits, 345-369. Cambridge: Cambridge University Press.
- Kramer, Ruth & Kyle Rawlins. 2011. "Polarity particles: An ellipsis account". *Proceedings of NELS 39*: 479-492.
- Krifka, Manfred. 2013. "Response particles as propositional anaphors". *Proceedings of SALT 23*: 1-18.
- Merchant, Jason. 2004. "Fragments and ellipsis". *Linguistics and Philosophy* 27 (6): 661-738.
- Pasquereau, Jeremy. 2018. "Responding to questions and assertions: embedded Polar Response Particles, ellipsis, and contrast." PhD diss., University of Massachusetts Amherst.
- Reich, Ingo. 2007. "Toward a uniform analysis of short answers and gapping". In *On information structure, meaning and form*, edited by Kerstin Schwabe and Susanne Winkler, 467-484. Amsterdam/Philadelphia: John Benjamins.
- Roberts, Craige. 2012/1996. "Information structure in discourse: Towards an integrated formal theory of pragmatics". *Semantics and Pragmatics* 5(6): 1-69.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. PhD diss., MIT.
- Ross, John Robert. 1969. "Guess who?". In *Proceedings from the Annual Meeting of the Chicago Linguistic Society 5*, edited by Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252-286. Chicago: University of Chicago, Chicago Linguistic Society.
- Sailor, Craig. 2012. "On Embedded Polar Replies". Paper presented at the Workshop on the Syntax of Answers to Polar Questions, Newcastle University 8-9 June 2012.
- Simons, Mandy. 2007. "Observations on embedding verbs, evidentiality and presupposition". *Lingua* 117: 1037-1056.
- Tanenhaus, Michael, and Greg Carlson. 1990. "Comprehension of Deep and Surface Anaphors". *Language and Cognitive Processes* 5(4): 257-280.
- Temmerman, Tanja. 2013. "The syntax of Dutch embedded fragment answers: on the PF-theory of islands and the WH/slucing correlation". *Natural Language and Linguistic Theory* 31(1): 235-285.
- Vicente, Luis. 2013. "In search of a missing clause." Paper presented at the *Parenthesis and Ellipsis workshop of the 34<sup>th</sup> Annual Meeting of the German Society of Linguistics*. Potsdam, March 2013.
- Weir, Andrew. 2014. *Fragments and clausal ellipsis*. PhD diss., University of Massachusetts Amherst.

Wurmbrand, Suzi. 2017. "Stripping and Topless Complements". *Linguistic Inquiry* 48 (2): 341-366.

## Notes

<sup>1</sup> For reasons of space, we provide hybrid gloss-translations (rather than interlinear glosses followed by natural English translations) for all non-English examples.

<sup>2</sup> We distinguish between the conjunction *și* 'and' and the homonymous form which behaves as an associative adverb.

<sup>3</sup> We distinguish between three different forms *nu* in Romanian: the proform negation *nu*<sub>1</sub> present in PVCs, the constituent negation *nu*<sub>2</sub> present in stripping, and the verbal negation *nu*<sub>3</sub> present in the verbal complex. Both proform *nu*<sub>1</sub> and constituent negation *nu*<sub>2</sub> have a lexical status, while the verbal negation *nu*<sub>3</sub> has an affixal status (see evidence in Barbu 2004).

<sup>4</sup> Both stripping and gapping constructions may occur in comparative contexts, but we leave aside these uses, since they are not relevant for the purpose of our study.

<sup>5</sup> An additional argument for distinguishing between stripping and gapping: unlike the gapped clause, the stripped clause can be an incidental adjunct (it may be mobile and have incidental prosody).

<sup>6</sup> We do not take into account here other alternative proposals, such as the (non-elliptical) across-the-board movement analysis, proposed by Johnson (2009, 2014).

<sup>7</sup> Weir (2014)'s terms are rather syntactic than semantic: non-factive verbs, such as *say*, *think*, *believe*, etc. are 'bridge' verbs, whereas factive verbs, such as *find out*, *confirm*, *be proud*, *be surprised*, etc. are 'non-bridge' verbs. For the purpose of our study, both semantic and syntactic terms cover the same empirical domain.

<sup>8</sup> Note that Romanian uses the same polarity particles *da* 'yes' and *nu* 'no' in both matrix and embedded contexts, unlike English (i.e. *yes* and *no* in matrix clauses and *so* and *not* in embedded configurations, cf. Sailor 2012).

<sup>9</sup> As anaphors, they must have an antecedent in the context. We distinguish between anaphors and catalepsis cases, the latter ones being illustrated in (i). We do not take into account these cataphoric uses in this paper (polarity particles in these cases could be considered as parentheticals which precede a full verbal clause).

- (i) **Constat că da**, omul este un animal social și este făcut să trăiască în societate, alături de alți oameni.  
'I notice that yes, man is a social animal and he is made to live in society, next to other people.'

## CHAPTER FOURTEEN

# THE ACQUISITION OF HUNGARIAN SUBJECTS IN A ROMANIAN-HUNGARIAN BILINGUAL CONTEXT

VERONICA TOMESCU

The paper charts the acquisition of Hungarian subjects in a Romanian-Hungarian bilingual context. The results show that the children are early on aware of the pragmatic constraints governing the null/overt subject alternation, as well as those governing the pre-/postverbal placement of the subject. However, in the earliest months, an overuse of overt subjects is observable, compared to the input and to L1 data, possibly as an effect of bilingualism. Since both the languages in question are null subject languages, it is argued here that no cross-linguistic interference could have occurred. The study also identifies sporadic word-order errors, due to the underspecification of the pragmatic constraints governing word order in Hungarian; aside from this being an interface phenomenon and thus vulnerable in bilingual acquisition, the errors can also be explained by cross-linguistic interference.

Keywords: Romanian-Hungarian bilinguals, null subjects, pronominal subjects, pre-/postverbal subjects.

### 1. Introduction

The null/overt subject alternation in null subject languages is a phenomenon at the syntax/discourse interface (Sorace and Filiaci 2006), hence it may be vulnerable in bilingual acquisition (Sorace et al 2009, a.o.). The present study documents the acquisition of Hungarian subjects in a Romanian-Hungarian setting, from the earliest utterances until the age of 9;0. Both Romanian and Hungarian are null subject languages, hence any quantitative differences compared to the input or to L1 can be expected to

have been caused by the effects of bilingualism itself and not to have been the effect of cross-linguistic influence.

Secondly, the existence of infelicitous postverbal subjects (focused or indefinite) is due to the underspecification of the strict pragmatic requirements governing word order in Hungarian under the influence of the less restrictive language, Romanian.

The paper is organized as follows. Section 2 describes Hungarian subjects and the similarities and differences between Romanian and Hungarian subjects, section 3.1 presents previous research on the acquisition of subjects in null subject languages, with particular emphasis on bilingual contexts, whereas subsection 3.2 describes the acquisition of subjects in Hungarian. Section 4 consists of the study proper.

## 2. Romanian and Hungarian subjects. Similarities and differences

Both Romanian and Hungarian are null subject languages. They allow null subjects in finite sentences and have no expletives (Avram and Coene 2008; É. Kiss 2004), see example (1a) for Romanian, and (1b) for Hungarian.

- (1) a. Ninge.  
      snows  
      b. Havazik.  
          snows  
          ‘‘It is snowing.’’

Overt personal pronoun subjects are felicitous especially when they bring new information to the discourse, for example when they are contrastively focused or indicate topic shift or topic continuity (Zafiu 2008, É. Kiss 1992, Kocsány 1995), and should be omitted otherwise.

Consider sentences (2), for Romanian. In (2a) the personal pronoun is contrastively focused. In (2b), the pronoun indicates primarily topic shift, although the topic continuity reading is also not excluded. An experimental study carried out by Teodorescu (2017) has found that adult speakers of Romanian accept both readings as equally plausible. The equivalent Hungarian pairs of sentences in (3) are very similar; the subject is contrastively focused in (3a) and conveys either topic shift or topic continuity in (3b). For similar contexts, Kocsány (1995) has found that if the pronoun is unstressed the topic shift reading is preferred (that is, the personal pronoun will be coindexed with the object of the preceding clause,

in our example Maria); conversely, if the stress falls on the pronoun the topic continuity reading is more acceptable – the personal pronoun will have as antecedent the subject, in our case Anna.

- (2) a. EL a venit, nu ea.  
 he has come not her  
 ‘It is he who has come, not her’  
 b. Ana a întrebat-o pe Maria; ea nu știe.  
 Ana has asked DOM Maria she not knows  
 ‘Ana asked Maria. She does not know.’
- (3) a. ŐK jöttek vissza.  
 they came back  
 ‘It is they who came back.’  
 b. Anna megkérdezte Máriát; ő nem tudja.  
 Anna asked Maria she not knows  
 ‘Ana asked Maria. She does not know.’

Additionally, Hungarian may require overt subjects in sentences with null copula. Since the 3<sup>rd</sup> person singular present tense copula is null, the presence of an overt subject (which can be a personal pronoun, like in 4a) may occasionally be necessary (É. Kiss 1992), for example when the predicative is a definite noun. See example (4a). However, as shown in (4b), an adjectival predicative can form the sentence on its own.

- (4) a. Ő a bajnok.  
 he/she the champion  
 ‘He/she is the champion.’  
 b. Ártatlan.  
 innocent  
 ‘He/she is innocent.’

Null subject languages typically allow both pre- and postverbal subjects, whose placement is governed by discourse requirements. However, these requirements are dissimilar in the two languages in the present study, as will be detailed below.

Hungarian focused constituents must necessarily move to the left of the verb (5a), to an operator position: the verb moves to the head of the Focus Phrase, with the focused constituent in its Specifier (É. Kiss 2004). However, the movement of the focused constituent is optional in Romanian (Alboiu 2002). Whereas the focused constituent may optionally move to

SpecIP (with the verb occupying the head of the Inflection Phrase, due to the verb-adjacency requirement), it is also allowed to remain in situ and the focus is signaled by means of phonological stress (5b).

- (5) a. ANNA jön, nem Mária.  
 Anna comes not Maria  
 b. ANA vine / Vine ANA, nu Maria.  
 Ana comes/ comes Ana not Maria  
 “It is Ana who is coming, not Maria.”

As shown in Alboiu (2002), in Romanian, preverbal subjects will usually indicate old/presupposed information, whereas new information is usually conveyed by means of presentational focus and should occur post-verbally. Furthermore, indefinite subjects may only occur preverbally under certain conditions (e.g. if they are contrastively focused, if they have partitive value, or if they are anchored by a postverbal locative phrase) (Alboiu 2002). Thus, in sentence (6), the indefinite subject, introducing new information, is preferred in postverbal position; it is permitted to occur preverbally only if it is balanced out – in our example – by the PP *the tree*. In example (6b), the preverbal subject has partitive value (one of the apples).

- (6) a. A căzut un măr./ (?)Un măr a căzut. /  
 has fallen an apple/ an apple has fallen /  
 Un măr a căzut din copac.  
 an apple has fallen from tree  
 “An apple has fallen down (from the tree).”  
 b. Un măr a căzut, ceilalți au rămas în copac.  
 “An apple has fallen, the others are still in the tree.”

In Hungarian, however, new information is often conveyed by means of identificational focus, which requires the focused element to appear to the left of the verb. Furthermore, indefinite (or bare) subjects generally function as verb modifiers and must move to the left of the verb (to a specifier of an Aspect Phrase in this case, whose head hosts the verb) (É.Kiss 2004). Thus the subject in (7) would not have been felicitous to the right of the verb.

- (7) Alma nagyságú jég esett<sup>1</sup>.  
 apple sized ice fell  
 “There fell hailstones the size of apples.”



By contrast, a similar sentence in Romanian, with an indefinite subject introducing new information, would be best with a postverbal subject – see (8). Both the Hungarian and Romanian examples are news headlines.

- (8) A căzut grindină cât oul de porumbel<sup>2</sup>.  
 has fallen hail as egg.DEF of pigeon  
 “There fell hailstones the size of pigeon’s eggs.”

And lastly, in Hungarian certain elements have an inherent focus feature and must always appear preverbally, for example nouns accompanied by *csak* ‘only’, *is* ‘too’ (É. Kiss 2004). Crucially, their Romanian counterparts need not move to the left periphery.

### 3. Previous research on the acquisition of subjects

#### 3.1 Previous research on the acquisition of subjects in bilingual contexts

The acquisition of subjects in bilingual contexts has been the focus of several well-known studies. The differences between bilinguals and monolinguals regarding the rate or interpretation of overt (pronominal) subjects have been explained as cross-linguistic influence (Müller and Hulk 2000): certain interpretable (syntactic/pragmatic) features in particular syntactic structures in one language may become underspecified and subsequently considered optional in case this same requirement is absent in the other language. More precisely, children acquiring a non-null subject language such as English in combination with a null subject language have been found to produce a higher rate of overt pronominal subjects in the null subject language than has been reported for monolingual children. This overuse may have been caused by the influence of the non-null subject language. Such results have been reported by Paradis and Navarro (2003) for Spanish in a Spanish-English context, by Haznedar (2007) for Turkish in a Turkish-English bilingual context. Similarly, in an experimental study, Serratrice (2007) found that Italian-English bilinguals performed differently from the monolingual Italian control group regarding the interpretation of overt pronominal subjects, most likely under the influence of English. Not all studies found evidence of cross-linguistic influence in bilingual development, however: Hinzelin (2003), in a German-Portuguese bilingual context; Juan-Garau and Perez-Vidal (2000) with an English-Catalan bilingual child; Zwanziger et al (2005) in an English-Inuktituk context; Serratrice (2002) with an English-Italian bilingual.

Crucially, the overuse of overt subjects has also been recorded even with combinations of two null subject languages, where cross-linguistic interference could not have played a part. Bilinguals often have difficulty with phenomena at the syntax/discourse interface independently of the language combination, and the null/overt subject alternation and the use and interpretation of anaphors are such an example. Thus, Sorace et al (2009) found that a group of Italian-Spanish bilinguals were more prone to accept overt subjects as felicitous than an age-matched group of monolinguals. Bonfieni (2018), in a study on pronoun interpretation by a group of Italian-Sardinian bilinguals, argues for the linguistic effects of bilingualism independent of cross-linguistic differences.

### 3.2 The acquisition of subjects in L1 Hungarian

Subjects emerge early in L1 Hungarian, before 2;0 (Balassa 1893, Wéber 2007, 2011). Both pre- and postverbal subjects are attested, alongside verbs with null subjects.

- (9) a. paci eles  
horse down-fall  
Intended: “the little horse fell down” (Lacika 1;7, Balassa 1893:66)
- b. ül Lacika  
sits Lacika  
“Lacika is sitting.” (Lacika 1;8, Balassa 1893:67)
- c. ül székbe  
sits chair-in  
“He is sitting on the chair.” (Lacika 1;8, Balassa 1893:67)

Pronominal subjects are attested early: for example, before the age of 2;0 in Wéber (2011), at 2;1 in the Andi CHILDES corpus (MacWhinney 2000), and at 2;2 in the Réger (2004) corpus. The pronominal subject in (10) is contrastively focused.

- (10) te vagy a buta  
you are the stupid  
“It’s you who are stupid.” (Miki, 2;2, Réger 2004)

Longitudinal studies on the acquisition of L1 Hungarian report varied results regarding word order preferences in the early stages. MacWhinney (1985) documents that the earliest word-order patterns attested are topic

before comment and focus before verb; the earliest sentences to be produced are however verb-initial – which is in fact different from the adult preference, since Hungarian sentences are mostly verb-final. Wéber (2007) reports of her own child that most utterances contain an Agent or a Patient following a verb or a particle. However, the child recorded in Dezsó (1970 in Wéber 2007) prefers subjects as topics (that is, preverbal subjects).

- (11) a. alszi(k) mama  
 sleeps mother  
 “Mother is sleeping.” (Jancsi 1;10, Wéber 2007: 224).
- b. Ági kakász  
 Ági poos  
 “Ági is pooing.” (L.J. 2;0, Dezsó 1970:90 in Wéber 2007: 225)

Subject-related word-order errors are practically non-existent in child Hungarian (MacWhinney 1985, Wéber 2007). As early as 1;10, children are aware of the correct word order in sentences with focus (Wéber 2007). For instance, see sentence (10) above, where the pronominal subject is contrastively focused: the copula (in Focus<sup>0</sup>) correctly follows the subject (in SpecFocus).

## 4. The study

### 4.1 Aim and method

The data used in the study comes from two sets of corpora.

The first set consists in two longitudinal corpora of spontaneous, naturalistic conversations<sup>3</sup>. The two target children (brothers) were recorded for 30-60 minutes a week, for a period of two years, and the two corpora overlap: additionally, the corpora contain utterances produced by their older brother who was present in most recordings. The three boys live in Bucharest with their Romanian monolingual father and Romanian-Hungarian bilingual mother. They are unbalanced bilinguals, with Romanian the dominant language. The children produce both Romanian and Hungarian utterances, according to preference and/or the identity of the interlocutors. For the present analysis, only Hungarian utterances were considered, containing a finite verb, with the exception of imperatives. Imitations, songs, etc. were excluded, as well as mixed Romanian-Hungarian utterances. The utterances were coded for null/overt subject, and the overt subjects were further coded for category and position (pre-

/postverbal). The age range and number of utterances taken into consideration for the present analysis are given in Table 1.

The second set of data consists in recordings collected several years later, from the same three brothers<sup>4</sup>. These corpora were collected in a more controlled setting: the children were explicitly requested to speak only in Hungarian, and initially encouraged to tell stories, but were not otherwise prompted. Mostly they ended up describing their drawings, but they also narrated (parts of) familiar stories or films; the stories are interspersed with some spontaneous utterances. In what follows, these corpora will be referred to as the ‘corpora of narratives’, in order to distinguish them from the longitudinal corpora. The age range is given in Table 1.

Child	Longitudinal corpora		Corpora of narratives	
	Age	Number of contexts	Age	Number of contexts
Toma	1;9-3;6	834	5;7-5;9	176
Petru	1;8-2;8	167 <sup>5</sup>	4;1-4;3	185
Matei	4;7-7;6	294	8;10-9;0	322

**Table 1. The data used in the analysis.**

Further, a number of 822 Hungarian utterances were selected for analysis out of the mother’s speech (from the longitudinal corpora).

The aim of the study was to chart the development of subjects in Hungarian; there are three research questions that an answer was sought for.

Firstly, the study looks into the emergence of subjects, focusing on the rate of overt subjects produced by the children in the earliest months. The analysis of the Romanian part of the longitudinal corpora in Tomescu (2018) found a higher rate of overt subjects than has been reported for L1 Romanian (Avram and Coene 2010, Teodorescu 2017); this is in line with previous research on the acquisition of subjects in bilingual contexts (Paradis and Navarro 2003, Haznedar 2007, etc.). Consequently, it was examined whether the two younger children in the study produce a higher rate of overt Hungarian subjects in the early months that is found in the input, that their older brother produces, and that is found in L1 Hungarian. To this latter end, I compared the results obtained from the longitudinal corpus to L1 Hungarian longitudinal data: the child Miki, in the Réger corpus (Réger 2004). The corpus (to be found on CHILDES, MacWhinney 2000) consists of 31 recordings between the ages 1;11 and 2;11.

Secondly, the study seeks evidence of any vulnerability in the use of pronominal subjects. Anaphors are notoriously difficult for bilinguals (Sorace et al 2009, Bonfieni 2018, Sorace 2018, etc.) who tend to overproduce them compared to their monolingual peers.

Thirdly, as presented in section 2, Hungarian focused subjects (contrastive, inherent or identificational focus) must move to the left periphery of the sentence; in Romanian focused constituents may remain in situ. Therefore, the data was examined for evidence that the children fail to front focused constituents in Hungarian.

Both phenomena, the use of anaphors and focus, are situated at the interface between syntax and pragmatics, therefore may be vulnerable in 2L1 acquisition (Sorace and Filiaci 2006, Sorace et al 2009, Sorace 2018, etc.). In order to answer these questions, the longitudinal corpora were not deemed sufficient, since it only covers an early stage of development, consequently the analysis was extended to the corpus of narratives.

The method of analysis for the corpus of narratives was as follows: all utterances with a finite verb and an overt subject were extracted (the number of utterances is given in Table 1). The utterances were coded for pronominal subjects, as well as pre-/postverbal subjects. Pronominal subjects were roughly divided into two categories: (1) pronouns which were focused or otherwise informative (contrastive topic or topic shift) and (2) pronouns which were uninformative and should have been omitted. Infelicitous postverbal subjects (which should have been preverbal in accordance with Hungarian discourse rules) were counted. The same method of analysis for pronominal and postverbal subjects was also applied for the longitudinal corpora.

The results are presented in the following section.

## 4.2 Results

### 4.2.1 Early subjects. The null/overt subject alternation

Subjects are attested quite early in Hungarian: at 1;9 in the Toma corpus and at 1;8 in the Petru corpus.

- (12) a. kapta Ma(tei)  
got Matei  
“Matei got (it).” (Toma 1;9)
- b. tata elment  
father away-went  
“Father has left.” (Petru 1;8)

- c. ír (T)oma  
writes Toma (Petru 1;8)

Hungarian pronominal subjects emerge at 1;11 in the Toma corpus (18), but are not attested at all in the Petru corpus (he does produce pronouns with other syntactic functions already at 1;7: the 3<sup>rd</sup> person singular dative *neki*, the second person singular *te*, etc.). Overall personal pronouns represent 11% of overt subjects in the Toma corpus. The pronoun in (14a) is contrastively focused. Other persons are represented in the later part of the corpus; see (14b), where the child employs the second person singular subject with a Topic shift interpretation.

- (13) a. ÉN ragasztom oda  
I stick-1SG there  
“I’ll stick it on myself.” (Toma 1;11)
- b. akarom menni be az árnyékba<sup>6</sup>. te szereted a napot?  
want go in the shade-in. you like the sun  
“I want to go into the shade. As for you, do you like the sun?”  
(Toma 3;3)

Contexts with null subjects outnumber contexts with overt subjects with all three children.

- (14) a. nem dobjuk ki  
not throw-1PL out  
“we are not throwing it out.” (Matei 4;7)
- b. doarme  
sleeps  
“[the bird] is sleeping.” (Toma 1;10)
- c. nu vreau acra  
not want-1SG sour  
“I don’t want a sour one.” (Petru 1;10)

In Matei’s utterances, we find a rate of 33% overt subjects overall, with no great variation over time.

In the Toma corpus, overt subjects represent 36% of all contexts overall. However, between 1;10-2;1, the rate of overt subjects in Hungarian is as much as 50%; it decreases after 2;2 and remains reasonably constant.

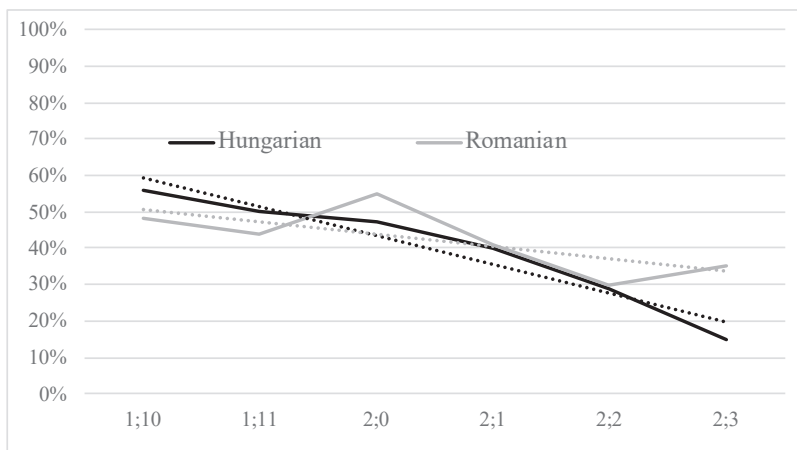
In the mother’s Hungarian utterances, we find a more or less constant rate of 39% overt subjects. The difference between CDS and Toma’s early

utterances (1;10-2;1) in overt subject rate is statistically significant:  $\chi^2 = 5.887$ ,  $df = 1$ ,  $p = 0.01$ .

I compared the rate of overt Hungarian subjects produced by Toma to the rate of overt subjects produced by the monolingual Hungarian child Miki, in the Réger corpus (Réger 2004). Miki's early utterances (1;11-2;1) contain a rate of 14% overt subjects; he produces overt subjects to a rate of 31% overall. The difference between Toma and Miki in the early months regarding overt subject production is statistically significant ( $\chi^2 = 8.815$ ,  $df = 1$ ,  $p = 0.003$ ). It can therefore be seen that in the early months the child produces a much higher rate of overt subjects both than in the input, and in the utterances of his older brother, and also compared to a monolingual Hungarian child.

As mentioned in the previous section, I showed in Tomescu (2018) that in the same period between 1;10 and 2;1, Toma's Romanian utterances also contain a significantly high rate of overt subjects: 47%. After the age of 2;2, the rate of overt Romanian subjects decreases in the Toma corpus and overall a percentage of 38% overt Romanian subjects was recorded for the period 1;10-2;11. These percentages were shown to differ significantly from the input—the mother produces a rate of 38% overt Romanian subjects. Furthermore, the percentages are also significantly higher than what has been reported for L1 Romanian (Avram and Coene 2010 and Teodorescu 2017).

To sum up, it appears that Toma's rate of overt subjects follows the same pattern in both languages: he starts out with a higher rate in the earlier months, which decreases towards the middle of his third year, see Figure 1.



**Fig. 1. Toma. Overt subjects.**

In the Petru corpus, overt subjects represent 43% of all contexts with finite verbs overall, between 1;8, when the first subjects are attested, and 2;3 (beginning with 2;4 only mixed utterances with finite verbs are attested). The tendency is increasing rather than decreasing. At 2;3 the child produces a rate of 60% overt subjects. A word of caution, however, the sample size is quite small for purposes of statistical comparison. Figure 2 represents the development path.

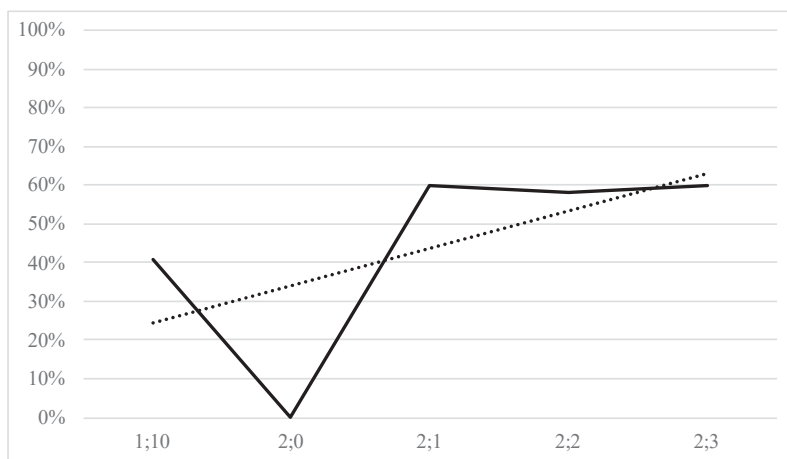


Fig. 2. Petru. Overt subjects. Hungarian.

#### 4.2.2 Word order errors

Only one word order error was identified, in Toma's utterances in the longitudinal corpus, with an infelicitous postverbal focused subject (15). In the corpus of narratives, 4 out of 65 (6%) subjects produced by Toma are incorrectly placed; in Petru's case, 6 out of 71 (8%). Below are some examples of infelicitous preverbal subjects. The inherently focused nouns 'the bat' and 'the ghost' in (16a) and the indefinite subject in (16b), which as new information counts as identificational focus.

- (15) akarom ÉN  
 want-1SG I  
 "I want to." (Toma 2;4)



- (16) a. \*és jött is a denevér, és jön is a szellem<sup>7</sup>  
 and came too the bat and comes too the ghost  
 Intended: “And then the bat came as well, and the ghost too.”  
 (Petru 4;1)
- b. ??és van ott egy lódarázs  
 and is there a hornet  
 Intended: “And there was a hornet there.” (Toma 5;7)

In Matei’s utterances in the two longitudinal corpora, only 4 postverbal focused personal pronoun subjects can be found, whereas focused subjects should surface preverbally. See for example the inherently focused subject in (17a).

- (17) a. \*akárom látni **én is**.  
 want see I too  
 “I want to see it too.” (Matei 4;10)

In Matei’s narratives, only 5 out of 160 (3%) overt subjects are incorrectly placed. In example (18), the subject (the earthworm) should have preceded the verb.

- (18) ??a cipőjén van **egy giliszta** cipőfűző helyett  
 the shoe-his-on is an earthworm lace instead  
 “There is an earthworm instead of laces on his shoe.” (Matei 8;10)

### 4.2.3 Overuse of pronominal subjects

In Toma’s utterances in the longitudinal corpus, therefore between the ages of 2;0 and 3;0, 19% of all pronominal subjects do not appear to be informative, do not seem to signal topic shift and are not contrastively focused. See examples (19a) and (19b), where no information is brought to the discourse by the use of the personal pronoun subject. The personal pronoun subject is not contrastively focused and has no topic shift interpretation, since the previous utterance has the same referent. By contrast, example (19c) is felicitous, with an inherently focused subject, as is example (19d), where the pronominal subject is contrastively focused.

- (19) a. nem akárom. ?**én** akárom. (Toma 2;11)  
 “I don’t want it. I want it.”

- b. MOTHER: akarod ezt a cinegét? “Do you want this chickadee?”  
 TOMA: ?**én** akarom ezt  
 I want this  
 “**I** want this one.” (Toma 2;10)
- c. és **én is** szeretem a mézet  
 and I too like the honey  
 “**I too** like honey.” (Toma 2;10)
- d. **ÉN** akarom a(z) ölébe ...  
**I** want the lap-her-in  
 Intended: “**I** want to sit in your lap.” (Toma 3;0)

However, the overuse of the first-person pronoun subject is not surprising in the utterances of a two-year-old whose view of the world is understandably egocentric. I showed in Tomescu (2018) that both Toma and Petru overused the first-person pronoun subject in Romanian: as much as 32% of Toma’s pronominal subjects and 13% of Petru’s do not seem to have any discourse function. See further examples from the Romanian utterances of the two longitudinal corpora in (20). The pronoun ‘eu’ need not have been used since it signals neither topic shift, nor focus. The infelicitousness is especially noticeable in Petru’s examples (20b), from the same recording, where he gives a running commentary of his activity: the purpose of the personal pronoun seems to be exclusively to draw attention to himself.

- (20) a. dacă sunt bebeluș ?**eu** am picioare mici  
 if am baby **I** have feet small  
 “if I am a baby **I** have small feet.” (Toma 2;9)
- b. desenez. ?**eu** desenez cu creta. ... ?**eu** desenez. ...  
 draw I draw with chalk.DEF I draw  
 ?**eu** urc cu Matei.  
 I climb with Matei  
 ‘I am drawing. **I** am drawing in chalk. **I** am drawing. I am climbing with Matei.’ (Petru 2;2)

Examples (21), from the Hungarian part of the Toma corpus, perhaps even better illustrate this tendency of the child to put himself forward in the discourse. The mother also uses overt personal pronouns with Topic shift interpretation, indicating to the child what he is allowed to do; the child promptly repeats the pronoun (in the correct person).

- (21) a. TOMA: és én is csinálom a Petru székét.  
 and I too make the Petru chair-of  
 “And I too am building Petru’s chair.”  
 MOTHER: **te** segítesz mamának.  
 you help mother ‘you are helping me’  
 TOMA: igen **én** segíték. **én** segíték  
 yes I help. I help (Toma 2;6)
- b. MOTHER: mama fogja és **te** bedugod  
 mother holds and you in-put  
 ‘Mother holds [the cannon] and you put [the  
 cannonball] in.’  
 TOMA: **én** bedugom  
 I in-put  
 “I put it in.” (Toma 2,6)

As for L1 Hungarian, the Réger corpus also contains similar less than felicitous utterances. In this corpus (Réger 2004), 31% of all overt subjects (n=208) are personal pronouns, 1<sup>st</sup> person singular. Of these, 30% are not obviously contrastively focused and do not signal topic shift; they seem superfluous and therefore could possibly have been omitted. See example (22).

- (22) de **én** kidobtam már. ?**én** kidobom a szemétbe .  
 but I out-threw already. I out-throw the garbage-in  
 “but I have thrown it away already. I am throwing it away.” (Miki 2;4, Réger 2004)

In the 2L1 longitudinal corpora, out of all overt subjects that Matei produces, 26% (n=25) are personal pronouns (one third person singular, the rest first person singular). He uses pronominal subjects with contrastive focus value (as in (23a, b)), or as Topics (23c). His examples are on the whole visibly better, the pronoun is more informative.

- (23) a. és ez az **ÉN** papucsom **ÉN** töltöm meg  
 and this the my slipper-my I fill up  
 “And this is *my* slipper so *I* be filling it up.” (Matei 5;8)
- b. de **Ő** nem akar.  
 but he not wants  
 “But *he* won’t (play).” (Matei 4;11)
- c. **én** azt hittem hogy az darázs.  
 I that thought that that wasp  
 “(As for me,) *I* thought it was a wasp.” (Matei 4;8)

In the corpus of narratives, there are very few overt personal pronoun subjects in the utterances of the oldest brother – 11 out of 160 overt subjects; none infelicitous: he correctly uses overt personal pronoun with topic shift or focus value. In example (24), the child explains that three animals have each certain knowledge and uses the personal pronoun subject to emphasise the identity of the respective character; in example (25), the overt singular and plural subjects, with topic shift value, are employed in an evident attempt to avoid ambiguity in narrating.

- (24) igen, a vakond, az tudta mert **ő** a föld alatt van és látta a sárgarépákat; a lepke azt tudja hogy melyik a káposzta mert **ő** káposztalepke; a csiga azt tudja hogy melyik<sup>8</sup> a saláták mert **ő** minden nap eszik salátát.

‘yes, the mole, that one knew because **he** is underground and saw the carrots; what the butterfly knows is which is the cabbage because **he** is a cabbage-butterfly; what the snail knows is which are the cabbages because **he** eats cabbage every day’

- (25) azt mondja nekik hogy ne mondja el a bátyjainak hogy hol van a gyűszű és aztán **ők** megmondják és áthagyják szélfiucskának és **ő** átadja a bátyjainak  
‘(he) tells them not to tell his brothers where the thimble is, and then **they** tell and give it to windboy and **he** gives it to his brothers’.  
(Matei 8;10)

In the corpus of narratives, Toma uses 9 personal pronoun subjects, out of 65 overt, 3 of which seem superfluous. In example (26), having finished his story whilst seated in front of a mirror, he is now describing to his mother and brother what his reflection is doing: the identity of the agent is obvious and need not have been explicitly mentioned.

- (26) (?)**én** \*integet<sup>9</sup> mamának; ?**én** \*integet Petrunak is mamának is  
I wave mother-to; I wave Petru-to too mother-to too  
Intended: “I am waving to mother; I am waving to both Petru and mother.” (Toma 5;9)

Note example (27) where Toma actually repeats the sentence adding an overt personal pronoun, in order to clarify and indicate that the subject of the second verb is not the same as the first.

- (27) és a szélfiúcska azt hitte hogy akar hintázn; azt hitte hogy ő akar hintázn  
 ‘and the windboy<sub>i</sub> thought that (he<sub>j</sub>) wants to swing; (he<sub>i</sub>) thought that he<sub>j</sub> wants to swing’ (Toma 5;7)

Petru however fares somewhat worse in the narrative corpus. He produces 23 personal pronouns out of 71 overt subjects, out of which 18 are uninformative. They are all 1<sup>st</sup> person singular pronouns. In fact, 22 of the 23 pronouns are 1<sup>st</sup> person singular, whereas Toma’s and Matei’s pronouns are more varied. Mostly he describes what he is pretending to be; the utterances in (28) are consecutive and the repetition of the pronoun is unnecessary. Additionally, the predicative should precede rather than follow the copula.

- (28) a. ??én vagyok egy teafőző  
 I am a teapot  
 b. ??én vagyok egy teahúzó  
 I am a tea-puller (?)  
 c. ??én vagyok egy kihúzó  
 I am a out-puller (?)  
 d. ??én vagyok egy foghúzó  
 I am a tooth-puller (Petru 4;1)

### 4.3 Discussion

The Interface Hypothesis (Sorace & Filiaci, 2006, Sorace 2011) predicts that structures at the interface between syntax and pragmatics will be acquired with some delay or difficulty compared to structures that only require syntactic computation. Hence, since in null subject languages the interpretation and overt realization of subjects is dependent on discourse factors, the Interface Hypothesis will predict vulnerability in bilingual acquisition. One explanation for this may reside in the fact that sentence-processing abilities are subject to cognitive control (Sorace 2011, 2018, Sorace and Serratrice 2009, Sorace et al 2009); evidence from psycholinguistic studies has shown that the two languages are constantly active while processing is carried out in one of the languages (Bialystok et al 2004, Luk et al 2011): consequently, a mechanism is needed for keeping the two languages separate, a mechanism which appears to use the same executive functions that are used to control focused and divided attention. On the one hand, this is an advantage, since bilinguals are consequently more efficient in tasks that require the suppression of distractors or

conflicting information (Luk et al 2011, Bialystok et al 2004, Bialystok et al 2008, Bialystok 2011), but the constant suppression of the unwanted language takes up cognitive resources needed for other tasks, namely, in our case, the use and interpretation of anaphors. The use of anaphors requires considerable cognitive resources, since the speaker must constantly update their awareness of the changes in the context and try to be aware of the cognitive environment they share with their interlocutor. And since bilinguals already use up some cognitive resources with the necessity to juggle between the two languages, they may find it difficult on occasion to use or interpret overt subjects in a fully target-like manner and end up overusing or misinterpreting pronominal subjects. Crucially, the differences between bilinguals and monolinguals with respect to anaphor use and interpretation are mainly quantitative and not qualitative, variation is observable within the same individual, and performance is always task-related (Sorace 2011, Sorace & Serratrice 2009, Sorace et al 2009). Alternatively, the overt form may be used as a default to minimize cognitive effort (Sorace 2018).

Another possible explanation is that bilinguals have a tendency to be 'overexplicit' (Sorace 2018), since they are more aware of potential ambiguities and are overly concerned with avoiding them. Bilinguals have enhanced Theory of Mind abilities (Kovács 2009, Goetz 2003): they have been shown to develop the cognitive abilities associated with Theory of Mind on average a year earlier than monolinguals. They also have superior metalinguistic abilities and are very early aware not only of their speaking two languages, but of the speakers around them speaking both, or only one of the two languages in question (Gawlitzeck-Maiwald & Tracy 2005, Genesee et al 1995, Serratrice et al 2009, Petitto et al 2001, Tomescu 2017), which makes them more aware of their potentially being misunderstood or not understood. Hence the tendency to overexplain and perhaps overuse overt forms to avoid ambiguity.

Therefore, it is not surprising that bilingual children in the present study use a higher rate of overt subjects in the earliest months, nor that they overuse personal pronouns.

Furthermore, monolingual children have also been found to overuse personal pronouns in the early years, as is the case of the monolingual Hungarian child Miki discussed here. A similar observation was made by Teodorescu (2017) regarding the subject production of a Romanian monolingual child, as well as by Serratrice (2005) in her study on the acquisition of Italian subjects. Admittedly, these monolingual children are younger than the bilingual children in the present study.

As for word order errors, they are practically non-existent in the early utterances of the two younger brothers, but are present in Toma's and Petru's narratives, when the children are older, as well as in Matei's utterances in the longitudinal corpora (but to a lesser extent in his narratives).

In Tomescu (2015) these same corpora of narratives are examined for word-order errors other than subject-related ones. The most frequent errors pertain to the postverbal placement of focused constituents, which is forbidden in Hungarian but permitted in Romanian, or the postverbal placement of bare/indefinite nouns, or oblique complements, which in Hungarian must be preverbal as noun modifiers (É. Kiss 2004), but in Romanian will usually occur in postverbal position, as informational focus, bringing new information to the discourse (Alboiu 2002).

The errors are evidently a result of cross-linguistic influence: as has been detailed in section 2, the requirements governing the pre- or postverbal position of the subject do not overlap in the two languages. Subjects which are obligatorily preverbal in Hungarian may or must be postverbal in Romanian.

Furthermore, since these word order requirements are at the syntax/discourse interface, they are predicted to be vulnerable in bilingual acquisition as an Interface phenomenon (Sorace and Filiaci 2006, Sorace 2011) and may come to be treated as optional. That the (two younger) children produce such word-order errors at a later age but not in their earliest utterances could perhaps be a result of their gradually speaking Hungarian less and less. Once they start kindergarten and spend an increased amount of time with their Romanian monolingual peers, their opportunities as well as desire to speak Hungarian will decrease. That interface phenomena, such as the discourse rules governing word order in sentences with focused constituents or the placement of new/old information in the sentence, should be the first to show vulnerability has also been found to be the case in L1 attrition (Tsimpli et al 2004): while purely syntactic phenomena are not vulnerable in acquisition and are not affected by attrition, phenomena at the syntax/pragmatics interface are the first to show signs of instability of optionality.

What happens is that the language where the constraints are more severe (Hungarian), where movement is obligatory, comes to be influenced by the less restrictive language (Romanian), where movement of the constituent is optional. This type of underspecification (Tsimpli et al 2004, Belletti et al 2007) has been attested in bilingual acquisition. Importantly, no wholesale transfer occurs: bilinguals only optionally and occasionally produce non-target structures (Serratrice 2013); they tend to have increased flexibility in

the use and interpretation of interface phenomena. Thus, the children also produce felicitous and target-like structures: the focused subjects are correctly moved to the left of the verb in (29a), (b), and (d), as are the inherently focused subjects in (29c) and (e).

- (29) a. most PETRU olvas  
now Petru (focused) reads  
“now it’s *Petru*’s turn to read.” (Toma 5;7)
- b. igen VÍZ van benne a bálnában  
yes water (focused) is inside the whale-in  
“yes, there is *water* inside the whale.” (Petru 4;0)
- c. én is akarom ülni mama mellett  
I too want-1SG sit-INF mother next-to  
“I want to sit next to mother too.” (Petru 4;1)
- d. ott a motor van.  
there the engine is  
“That’s the engine in there.” (Matei 4;10)
- e. a tej is mocskos lett  
the milk too dirty became  
“The milk became dirty too.” (Matei 8;11)

## 5. Conclusion

To sum up, the paper examined three aspects in the acquisition of Hungarian subjects: the null/overt subject ratio compared to the input or L1 Hungarian, the felicitous use of overt pronominal subjects and the possible existence of word order errors due to the differences in the discourse rules governing the distribution of focused and indefinite subjects in the two languages. In the early months, the two younger bilinguals produce a significantly higher number of overt subjects than a Hungarian monolingual child (Réger 2004), than their older brother, and that there are in the input. The occasional overuse of pronominal subjects is noticeable throughout the corpora. The overuse of overt and pronominal subjects is not a case of cross-linguistic influence, since Romanian is also a null-subject language. Thirdly, the children occasionally produce word-order errors in the placement of focused subjects or subjects introducing new information to the discourse, but in this case as a result of cross-linguistic influence: whereas in Hungarian such subjects should occur preverbally, the children seem to treat this rule as optional, following the less restrictive word-order pattern in Romanian. All three phenomena exhibit vulnerability which has



been explained as an effect of bilingualism, vulnerability at the discourse/syntax interface (Sorace and Filiaci 2006, Sorace 2011).

## References

- Alboiu, Gabriela. 2002. *The Features of Movement in Romanian*. Bucharest: Editura Universității din București.
- Avram, Larisa, and Martine Coene. 2008. “Can children tell us anything we did not know about parameter clustering?”. In *The Limits of Syntactic Variation*, edited by Theresa Biberauer, 459-482. Amsterdam: John Benjamins.
- Avram, Larisa, and Martine Coene. 2010. “Early subjects in a null subject language”. In *New Directions in Language Acquisition: Romance Languages in the Generative Perspective*, edited by Pedro Guijarro-Fuentes, and Laura Dominguez, 189-217. Newcastle: Cambridge Scholars Publishing.
- Balassa, József. 1893. “A gyermek nyelvének fejlődése”. *Nyelvtudományi Közlemények* 23: 60-73, 129-144.
- Belletti, Adriana, Elisa Bennati, Antonella Sorace 2007. “Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian”. *Natural Language and Linguistic Theory* 25 (4): 657-689.
- Bialystok Ellen, Fergus Craik, Raymond Klein, Mythili Viswanathan 2004. “Bilingualism, Aging, and Cognitive Control: Evidence from the Simon Task”. *Psychology and Aging* 19 (2): 290–303.
- Bialystok, Ellen 2011. “Coordination of executive functions in monolingual and bilingual children”. *Journal of Experimental Child Psychology* 110 (3): 461–468.
- Bialystok, Ellen, Gigi Luk, Fergus Craik. 2008. “Executive control and lexical access in younger and older bilinguals”. *Journal of Experimental Psychology* 34 (4): 859-873.
- Bonfieni, Michela. 2018. “The Bilingual Continuum: Mutual Effects of Language and Cognition”. PhD diss., University of Edinburgh.
- É. Kiss, Katalin. 1992. “Az egyszerű mondat szerkezete.” In *Strukturális magyar nyelvtan. 1. Mondattan*, edited by Ferenc Kiefer, 79-177. Budapest: Akadémiai.
- É. Kiss, Katalin. 2004. *The Syntax of Hungarian*. Cambridge: Cambridge University Press.
- Gawlitzek-Maiwald, Ira, Rosemarie Tracy, 2005. “All Children Start out as Multilinguals”. In *Proceedings of the 4th International Symposium on*

- Bilingualism*, edited by James Cohen, Kara T. McAlister, Kellie Rolstad, and Jeff MacSwan, 875-889. Somerville MA: Cascadilla Press.
- Genesee, Fred, Elena Nicoladis, Johanne Paradis. 1995. "Language differentiation in early bilingual development". *Journal of Child Language* 22 (3): 611-631
- Goetz, Peggy J. 2003. "The effects of bilingualism on theory of mind development". *Bilingualism: Language and Cognition* 6 (1): 1-15.
- Haznedar, Belma. 2007. "Crosslinguistic Influence in Turkish-English Bilingual First Language Acquisition: The Overuse of Subjects in Turkish". In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition, North America*, edited by Alyona Belikova, Luisa Meroni and Mari Umeda, 124-134. Somerville, MA: Cascadilla Proceedings Project.
- Hinzelin, Marc-Olivier. 2003. "The acquisition of subjects in bilingual children. Pronoun use in Portuguese-German children". In *(In)Vulnerable Domains in Multilingualism*, edited by Natasha Müller, 107-137. Amsterdam: Benjamins.
- Juan-Garau, Maria, and Carmen Perez-Vidal. 2000. "Subject realization in the syntactic development of a bilingual child". *Bilingualism: Language and Cognition* 3 (3): 173-19.
- Kocsány, Piroska. 1995. "Műhelytanulmány az ő névmásról." *Magyar nyelvőr* 119: 285-293.
- Kovács, Agnes Melinda 2009. "Early bilingualism enhances mechanisms of false-belief reasoning". *Developmental Science* 12 (1): 48-54.
- Luk, Gigi, Eric de Sa, Ellen Bialystok. 2011. "Is there a relation between onset age of bilingualism and enhancement of cognitive control?". *Bilingualism: Language and Cognition* 14 (4): 588-595.
- MacWhinney, Brian. 1985. "Hungarian language acquisition as an exemplification of a general model of grammatical development". In *The crosslinguistic study of language acquisition, Vol.2 Theoretical issues*, edited by Dan Isaac Slobin, 1069-1155. Hillsdale NJ: Lawrence Erlbaum Associates.
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Müller, Natasha, and Aafke Hulk. 2000. "Cross-linguistic influence in bilingual children: object omissions and Root Infinitives". In *Proceedings of the 24th Annual Boston University Conference on Language Development*, edited by S. Catherine Howell, Sarah A. Fish, Thea Keith-Lucas, 546-557. Somerville, MA: Cascadilla, Press.
- Paradis, Johanne, and Samuel Navarro. 2003. "Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and

- English: what is the role of the input?”. *Journal of Child Language* 30: 371-393.
- Petitto, Laura Ann, Marina Katerelos, Bronna G. Levy, Kristine Gauna, Karine Tetreault, Vittoria Ferraro 2001. “Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition”. *Journal of Child Language* 28: 453-496.
- Réger, Zita. 2004. TalkBank.
- Serratrice, Ludovica. 2002. “Overt subjects in English: evidence for the marking of person in an English-Italian bilingual child”. *Journal of Child Language* 29 (2): 327-355.
- Serratrice, Ludovica. 2005. “The role of discourse pragmatics in the acquisition of subjects in Italian”. *Applied Psycholinguistics* 26 (3): 437-462.
- Serratrice, Ludovica. 2007. “Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English–Italian bilingual children”. *Bilingualism: Language and Cognition* 10: 225-238.
- Serratrice, Ludovica 2013. “Cross-linguistic influence in bilingual development. Determinants and mechanisms”. *Linguistic Approaches to Bilingualism*, 3 (1): 3-25.
- Serratrice, Ludovica, Antonella Sorace, Francesca Filiaci, Michela Baldo. 2009. “Bilingual Children's Sensitivity to Specificity and Genericity: Evidence from metalinguistic awareness”. *Bilingualism: Language and Cognition* 11 (2): 239-257.47.
- Sorace, Antonella. 2011. “Pinning down the concept of ‘interface’ in bilingualism”. *Linguistic Approaches to Bilingualism* 1 (1): 1-33.
- Sorace, Antonella. 2018. “Pronominal reference in bilingual speakers of Romance languages (or: let’s not miss the forest for the trees)”. Paper presented at *Romance Turn 9*, Bucharest.
- Sorace, Antonella, and Francesca Filiaci. 2006. “Anaphora resolution in near-native speakers of Italian”. *Second Language Research* 22 (3): 339-368.
- Sorace, Antonella, and Ludovica Serratrice. 2009. “Internal and external interfaces in bilingual language development: beyond structural overlap”. *International Journal of Bilingualism* 13 (2): 195-210.
- Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci, Michela Baldo. 2009. “Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children”. *Lingua* 119: 460-477.
- Teodorescu, Otilia. 2017. “The Acquisition of the subject in child Romanian at the syntax-discourse interface”. PhD diss, University of Bucharest.

- Tomescu, Veronica. 2013. *The syntax and acquisition of particles in English, Hungarian and Romanian*. Bucharest: Editura Universității din București.
- Tomescu, Veronica. 2015. "Word order errors at the discourse-syntax interface with Romanian-Hungarian bilinguals". *Revue Roumaine de Linguistique*, LX (1): 29-45.
- Tomescu, Veronica. 2017. *Acquisition in a Romanian-Hungarian bilingual setting*. Bucharest: Editura Universității din București.
- Tomescu, Veronica. 2018. "The acquisition of Subjects in 2L1 Romanian". *Bucharest Working Papers in Linguistics XX* (1): 39-63.
- Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock, Francesca Filiaci. 2004. "First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English". *International Journal of Bilingualism* 8 (3): 257-277.
- Wéber, Katalin. 2007. "Elemi tárgyias predikátumok jelentésszerkezetének kiépülése a magyarban (L1, L2)". In *Alknyelvdok Proceedings*, edited by Tamás Váradi, 218-232. Budapest: Nyelvtudományi Intézet.
- Wéber, Katalin. 2011. "Rejtelmes kétféleség. A kétféle igeragozás elkülönülése a magyar nyelvben. Megkülönböztetésük a magyar nyelvészeti hagyományban és gyermeknyelvi megnyilatkozások longitudinális korpuszvizsgálata alapján". PhD diss, University of Pécs.
- Zafiu, Rodica. 2008. "Anafora". In *Gramatica limbii române, II. Enunțul*, edited by Valeria Guțu Romalo, 749-766. Bucharest: Editura Academiei Române.
- Zwanziger, Elizabeth, Shanley Allen, and Fred Genesee. 2005. "Crosslinguistic influence in bilingual acquisition: subject omission in learners of Inuktitut and English". *Journal of Child Language* 32 (4): 893-909.

## Notes

---

<sup>1</sup> <https://www.agrotrend.hu/hireink/nem-lazsalnak-a-jegkarmerszeklo-rendszerberendezesei>.

<sup>2</sup> <https://www.opiniatimisoarei.ro/a-cazut-grindina-cat-oul-de-porumbel-nu-departe-de-timisoara/05/05/2018>.

<sup>3</sup> First described in Tomescu (2013).

<sup>4</sup> First described in Tomescu (2015).

<sup>5</sup> Many of Petru's Hungarian utterances do not contain finite verbs. Additionally, many of Petru's Hungarian utterances contain instances of switching and were thus not included in this analysis. Consequently, the number of utterances suitable for the present analysis is considerably smaller than in his brother's case.

<sup>66</sup> Incorrect word order: the particle should have raised to the left of the modal-like finite verb: *be akarok menni*; alternatively, in fact, the Location should have

---

functioned as verb modifier, replaced the particle and appeared in sentence initial position, to the left of the finite verb.

<sup>7</sup> Note an additional word order error: the particle *is* ‘too’ should follow rather than precede the noun.

<sup>8</sup> Correct: *melyek*.

<sup>9</sup> 1<sup>st</sup> person subject agreement missing.

## CHAPTER FIFTEEN

# 5-YEAR-OLDS ARE PRECISE WITH CARDINALS: EXPERIMENTAL EVIDENCE FROM ROMANIAN CHILD LANGUAGE

ADINA CAMELIA BLEOTU

On the basis of experimental evidence, the current paper argues that 5-year-old Romanian monolingual children have a precise, exact understanding of cardinal numbers. The finding can be accounted for either within a lexical framework, according to which cardinals have a precise meaning, or within an implicature-based framework, according to which children are able to derive implicatures with cardinals from early on, earlier than with quantifiers. The latter view is supported by children's frequent exposure to cardinals (through counting), the conceptual simplicity of cardinal scales as opposed to quantificational scales a.o.

**Keywords:** Romanian L1, cardinal numbers, scalar implicatures, precise semantics

### 1. Aim

The aim of this paper is to examine whether Romanian 5-year-olds derive scalar implicatures with cardinal numbers and to account for any possible differences between children and adults in this respect. On the basis of experimental evidence from language acquisition, the paper essentially argues that Romanian children (Age 5) are precise with cardinal numbers, but that several accounts are compatible with this fact: (i) a lexical account where children's behaviour is due to the fact that cardinal numbers are lexically stored with a precise, exact meaning, or (ii) an implicature-based account where children's precision is due to their ability to derive implicatures with cardinal numbers from early on. While children have been shown to derive less implicatures with existential

quantifiers than adults, evidence from experiments into the acquisition of cardinals seems to show that children perform much better with cardinal numbers (Papafragou & Musolino 2003, Hurewitz et al. 2006, Huang, Spelke & Snedeker 2013). This has led to a differential treatment of various scales depending on the items at stake. Previewing the results, the current experiment reveals no significant differences between children and adults in the treatment of cardinal numbers in Romanian, thus supporting a view where cardinals represent a different kind of scale than quantifiers. It remains an open question whether this is because cardinals are lexical encoders of exactness or simply able to trigger implicatures more easily due to various factors (e.g. children's early and frequent exposure to cardinals, and, especially, the importance of counting in early child education, which might help children associate cardinals with precise numbers, the hypothesis that cardinal scales might be conceptually simpler than quantificational scales, which involve handling set selection/inclusion, the importance of complex number systems in various languages a.o.)

The roadmap of the paper is as follows: I start by presenting previous research in the domain of scalar implicatures, focusing on cardinals (Section 2), then I move on to the presentation of the current experiment I conducted (*Hypothesis, Participants, Procedure, Materials, Results*) (Section 3), and then I discuss the accounts compatible with the experimental data.

## 2. Introduction. Research in the domain of scalar implicatures (with cardinals)

Scalar implicatures represent inferences drawn in conversation by means of violating the maxims of communication (Grice 1989). To take a simple example, a sentence such as (1):

(1) Some kitties are orange.

may be interpreted (and is usually interpreted by adults) as implicating that *some, but not all* kitties are orange. However, given that implicatures have the property of cancellability (Grice 1989), they can be explicitly contradicted by the speaker without the result being felt to be contradictory.

(2) *Some kitties. . . and maybe all of them* are orange.

According to Grice's scheme, in producing (1), in case the speaker is aware that all the kitties he/ she is talking about are orange, then he/ she has violated the maxim of informativeness (or Quantity):

(3) **Quantity maxim**

- i. Make your contribution as informative as is required.
- ii. Do not make your contribution more informative than is required.

More specifically, the speaker has violated the first submaxim, choosing a weak term from *a scale (Horn scale)*, i.e. a range of items ordered in terms of informational strength (in this case, the scale *<all, ..., some>*) (Horn 1972). Hence, ideally, the speaker should have said:

(4) All kitties are orange.

When confronted with the challenge of interpreting an utterance like (1), most adults generate scalar implicatures. However, children do not interpret underinformative sentences with scalar terms in the same way as adults. Research in the domain of the acquisition of scalar implicatures has revealed that children opt for a more logical interpretation instead of a pragmatic one, interpreting *some* as compatible with *all* (*some, possibly all*) (Noveck 2001, Papafragou & Musolino 2003, Pouscoulous et al. 2007, Katsos & Bishop 2011, Stoicescu, Sevcenco & Avram 2015 a.o.). In other words, in a context where, for instance, all kitties are in the teacup, when children are asked whether they agree with utterance (5), most children will say yes, unlike adults, who will mostly reject the utterance, assigning to *some* the pragmatic interpretation *some, not all*.



**Fig. 1: Kitties in a teacup**



## (5) Some kitties are in the teacup.

However, the picture is not the same for all types of scales: notably, there seem to be differences in whether or to what extent implicatures are generated with certain scales versus others both for children and adults. For this reason, it is important to distinguish between various kinds of scales:

- (i) quantificational, which may involve:
  - a) existential quantifiers (e.g. *some*-< *all*, *some*> scale)
  - b) modal verbs, adverbs, adjectives (e.g. epistemic *might*-< *must*, *might*> scale)
- (ii) numerical-with cardinal numbers (e.g. *two*-< *three*, *two*> scale)
- (iii) lexical-with lexical items (e.g. *pretty*-< *pretty*, *beautiful*> scale)

The question of particular interest to the current paper is how adults and children interpret sentences with cardinal numbers. When confronted with a sentence like (6), do adults and/or children usually understand *two* as meaning “not three” or do they interpret *two* as meaning “two...possibly three”?

## (6) Ann has two children. -&gt; NOT three children

Importantly, acquisition studies (Papafragou & Musolino 2003) seem to show that, just as in the case of existential quantifiers, children derive less scalar implicatures with cardinal numbers than adults.

However, there seems to be a significant difference between quantificational and numerical scales: children do not treat all scales alike, being more adult-like (though not fully) in their understanding of the scale of cardinal numerals (*two*< *three*) than in their understanding of the quantifier scale (*some*< *all*) (Papafragou & Musolino 2003).

Papafragou & Musolino’s (2003) work is an important study on implicatures with various scale types, and, given that the current paper draws a lot on the experiments conducted by the authors and the observations they make, it is useful to present them in more detail. Papafragou & Musolino (2003) investigated whether Greek children and adults derive implicatures with various scales and to what extent.

*Experiment 1* was conducted on 30 5-year-olds and 30 adults (all native speakers of Greek) on three different scales, <*oli*, *meriki*> (<all, some>), <*tris*, *diol*> (<three, two>) and <*teliono*, *arxizo*> (<finish, start>). Subjects watched certain characters (toy animals/ people) perform certain

activities in front of them. These performances represented contexts which satisfied the semantic content of the stronger (i.e. more informative) terms on each scale (i.e. *all*, *three* and *finish*) but were described using the weaker terms of the scales (i.e. *some*, *two*, *start*):

- (7) Merika apo ta aloga pidiksan pano apo to fraxti.  
Some of the horses jumped over of the fence.  
'Some of the horses jumped over the fence.'  
*Context*: All of the horses jumped over the fence.
- (8) Dio apo ta aloga pidiksan pano apo to fraxti.  
Two of the horses jumped over of the fence.  
'Two of the horses jumped over the fence.'  
*Context*: Three horses jumped over the fence.

A puppet described the situation by means of an utterance, and subjects were asked whether the puppet 'answered well' (i.e. *Apantise kala*, 'Did-(she)-answer well?'). This represents an important modification to the original TVJT (Crain & Thornton 1998), where subjects were asked if the puppet was 'right' or 'wrong'. It was felt that phrasing the question in terms of right/ wrong would result in acceptance of underinformative sentences, which are true but not pragmatically optimal, whereas phrasing the question in terms of 'answering well or not' would be more in tune with the idea of pragmatic adequacy. This modification was meant to encourage the production of more implicatures, and, hence, the rejection of underinformative sentences.

The results show that, unlike adults, who rejected these infelicitous descriptions to a high degree, children almost never did so. However, children's rejection rate on the numerical scale was reliably higher than on the two other scales, although not fully adult-like.

In *Experiment 2*, children were trained to detect pragmatically inadequate statements produced by a "silly puppet" e.g. children were encouraged to consider the statement *This is a small animal with four legs* "silly" in comparison to *This is a dog*. Moreover, the experimental set-up was modified to focus on a character's performance in a task. For example, in one of the stories, Mr. Tough brought back three horses; when asked how Mr. Tough did, the puppet answered *He caught some/two of the horses*. These modifications were meant to lead to more implicatures. Indeed, the results reveal that 5-year-olds were more likely to compute scalar implicatures, even though still not at adult-like levels. Once again, children fared better with numerical scales than quantificational scales.

In a different study by Hurewitz et al. (2006), 3- and 4-year-old English children from the US were asked to find pictures in which *The alligator took some of the cookies*, as well pictures in which *The alligator took two of the cookies*. Importantly, when faced with a picture in which the alligator took some of the cookies and a picture in which the alligator took all of the cookies, children selected both pictures. However, when faced with a picture in which the alligator took two of the cookies and a picture in which the alligator took four of the cookies, children selected only the picture in which the character had exactly 2 cookies out of 4, rejecting the one in which he had all 4. This suggests that cardinals have a different acquisition path than scalar quantifiers.

Similar findings were reported in Huang, Spelke & Snedeker (2013), who tested both adults and 2- and 3-year-olds in a covered box task meant to tease apart semantic and pragmatic aspects of interpretation. In the case of cardinal numbers, participants were asked *Give me the box with two fish* in the context of a visible mismatch (a box with 1 fish), a visible and salient lower bounded target (a box with 3 or 5 fish) and a covered box with unknown contents. The authors' purpose was to create a context where scalar implicatures are suspended, and this was done by embedding the cardinal within a singular definite NP (*the box with two fish*), leading to a unique referent presupposition in the context that satisfies the description (Frege 1892, Strawson 1950). In this way, generating implicatures would bring no additional information. Thus, Huang, Spelke & Snedeker (2013) predicted that, if *two* has lower-bounded semantics (*at least two*), then the choice should be the box with 3 or 5 fish. However, if *two* has an exact semantics (*exactly two*), since this option is not available, subjects should infer that the covered box must be the one with 2 fish. Importantly, both adults and children consistently gave exact interpretations for number words, picking the covered box. Children's tackling of cardinals is also significantly different from their behaviour with *some*. In a similar experiment with *some*, when asked *Give me the box where Cookie Monster has some of the cookies*, children chose the covered box significantly less than in the cardinal case. Huang, Spelke & Snedeker (2013) take the results to show that children derive implicatures with existential quantifiers, but that exact readings of cardinals are not derived via implicatures, but rather lexically available from the onset of acquisition. However, it could be that children have strengthened the mechanism of deriving implicatures to a great extent, due to their exposure or engagement in a variety of activities involving numbers, for instance. If that is the case, then the reason for children's preference for the covered box could simply reflect their general preference for

strengthened implicatures. In other words, it is not immediately clear that cardinals are stored with exact meanings in children's lexicon.

### **3. The Current Experiment**

The current experiment sets out to see how Romanian 5-year-old children interpret cardinal numbers in comparison to adults, whether they (seem to) derive scalar implicatures or not.

#### **3.1 Hypotheses**

The expectation is that Romanian children will behave similarly to the Greek and English children previously tested for implicatures, namely, they will provide an exact interpretation of cardinals. Given that previous studies show either significant or non-significant differences among children and adults, it is unclear whether children will fully pattern like adults or not.

#### **3.2 Participants**

23 children from a kindergarten in Bucharest, Romania (age range: 4;5-6;3, mean age: 5;45) took part in the experiment. In addition, the control group involved 23 adults (age range: 18-22, mean age: 21; 5) recruited from undergraduates at the Faculty of Foreign Languages and Literatures, at University of Bucharest.

#### **3.3 Procedure**

The method used was an adapted truth value judgment task (TVJT) (Crain & Thornton 1998). Children were introduced to a puppet who asked them to help a country girl describe the states and activities of various animals on her farm, as she was a bit confused and she did not want to upset her mother by not saying things well. Subjects were shown pictures depicting domestic animals and were asked whether they agreed with certain statements made by the girl about the pictures or not and why. Questions such as whether a certain sentence was right or wrong were avoided.

### 3.4 Materials

Subjects were given 12 randomized sentences: (i) 6 test sentences, and (ii) 6 control sentences, along with 12 pictures on the basis of which they were supposed to evaluate these sentences.

The 6 test sentences were all underinformative true sentences containing cardinal numerals, in which a certain property or event was predicated about a smaller number of animals than the pictures showed:

- (9) Doi câini au limba scoasă.  
two dogs have tongue out.  
'Two dogs have their tongue out.'
- (10) Trei pui stau în picioare.  
three chicks stand in legs  
'Three chicks are standing (up).'



**Fig. 2: Picture with dogs**



**Fig. 3: Picture with chicks**

For a full list of the test sentences, see **Table 1**, for a full list of the pictures used, see **Appendix**.

The 6 control statements included 3 statements with *toate* “all” (2 true statements and 1 false one), and 3 statements with *unii* “some” (2 false statements and 1 true statement), none of which was underinformative.

False sentences were included in the experiment in order to control for a possible yes bias in the case of children.

### 3.5 Results

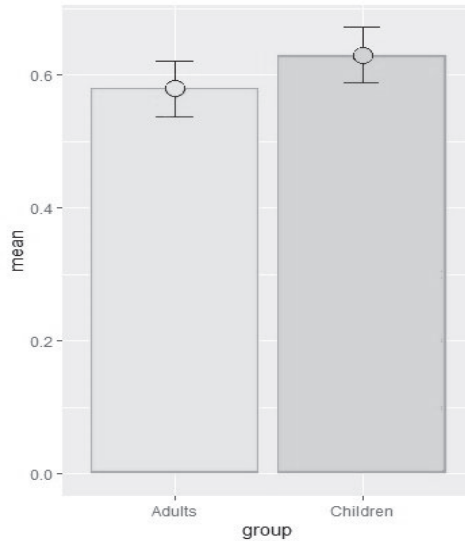
Both children and adults performed almost at ceiling in the control items, while their performance was less accurate in the underinformative contexts. Most of the children (63.04%) gave exact rather than logical answers, behaving very similarly to adults (57.97%)- see Table 1 and Figure 4.

Just like adults, most of the children were consistent in their answers, giving the same kind of answer for all underinformative sentences. Out of the 23 children, only 4 kids displayed variable behavior: 2 children gave 3 *yes* answers and 3 *no* answers, 2 children gave 4 *yes* answers, 4 *no* answers. This suggests the existence of a stage where exact/pragmatic and logical reasoning are in competition (at least in the case of some children). Some adult subjects also displayed patterns of thinking that showed competition between exact/pragmatic and logical thinking: there were 3 cases of adults with variable behaviour. Nevertheless, the majority of adults gave consistent answers.

In the **statistical analysis**, I used R (R Core Team 2015) and *lme4* (Bates et al. 2015) to perform a logistic regression with Answer as a dependent variable, Age Group and Sentence type (underinformative or control) as fixed effects and subjects and items as random effects. There was an effect of Sentence type ( $\beta=-7.047$ ,  $Z= 1.158$ ,  $p<0.01$ ), but no main effect of Age Group ( $\beta=0.004$ ,  $Z= 0.058$ ,  $p=0.457> 0.01$ ) and no interaction between Age Group and Sentence Type ( $\beta=-0.025$ ,  $Z= 0.051$ ,  $p=0.617>0.01$ ). It thus seems to be the case that children treat cardinal numbers adult-like from early on, possibly deriving scalar implicatures of the type “two, not three”, “three, not four”.

Test sentences	Number of exact answers given by children	Number of exact answers given by adults
<i>Doi cai au pete albe.</i> 'Two horses have white spots.'	14	13
<i>Doi câini au limba scoasă.</i> 'Two dogs have their tongue out.'	14	14
<i>Trei pui stau în picioare.</i> 'Three chicks are standing up.'	13	13
<i>După gard se află doi măgari.</i> 'Behind the fence there are two donkeys.'	18	14
<i>În pat dorm trei pisici.</i> 'Three kitties are sleeping on the bed.'	13	13
<i>Pe pajiște trei vaci mănâncă iarbă.</i> 'Three cows are eating grass on the pasture.'	15	13
Total	87 (Mean: 14.5 Proportion: 63.04%)	80 (Mean: 13.3, Proportion: 57.97%)

**Table 1. Test sentences and number of exact answers per test sentence**



**Fig. 4: Mean of exact answers with cardinals: Comparison between children and adults**

#### 4. Account

While Stoicescu, Sevcenco & Avram (2015) previously found a significant difference between Romanian children and adults in the interpretation of underinformative sentences with existential quantifiers, no such difference has been detected in the current experiment. It seems to be the case that children treat the numerical scale in a different way than the quantificational scale.

The results are different from those obtained by Papafragou & Musolino (2003), where there was a significant difference between children and adults even in the case of the numerical scale. In the current experiment, children and adults showed a very similar behaviour. These results are compatible with two possible accounts: (i) a lexical account where the meaning of cardinal numbers is precise, exact to begin with, (ii) a pragmatic account where cardinal numbers are interpreted as precise because of the generation of scalar implicatures.



#### 4.1 Notes on the Semantics of Cardinal Numbers

In order to evaluate what the best account for the data might be, let us delve into the semantics of cardinal numbers. Various linguists (Horn 1972, Gazdar 1979, Levinson 1983, 2000, Kadmon 1987, Rothstein 2017) have tried to tackle the issue of the meaning of cardinal numbers, coming up with different proposals.

A fact that has been repeatedly noticed in the literature is that numerical cardinals can be interpreted as meaning ‘exactly *n*’ (in which case there is an upper bound) or ‘at least *n*’ (in which case there is no upper bound), and that different contexts favour each reading. For instance, taking a look at these examples Rothstein (2017:21), one can notice that contexts (11a, b) favour an *exact* reading, whereas context (11c) favours an *at least* reading:

- (11) a. I have four dogs.  
 b. A dog has four legs.  
 c. A person with four dogs gets a reduction at the animal clinic.

In (11a, b), the meaning of *four* is *exactly four*, whereas in (11c), it is *at least four*. Any theory of the meaning of cardinal numbers has to be able to account for the existence of these two readings of cardinal numbers.

Importantly, one fundamental question addressed by semantic theories of cardinality is which of the two readings is the basic reading (Rothstein 2017). According to Horn (1972), the *at least n* reading is basic, and the *exactly n* reading is derived via scalar implicatures. The assumption in this case is that the speaker is observing Grice’s Maxim of Quantity and being as informative as possible. In other words, the speaker gives as much info as he/she should. This represents a pragmatic account of exact readings of cardinals. In contrast, another view (embraced by Landman 2003, following observations in Kadmon 1987, Kamp & Reyle 1993, Ariel 2006, Cohen & Krifka 2011), holds that the *exactly n* reading is basic, and the *at least n* reading is derived from it via existential quantification. Evidence in favour of such a view comes from the fact that the *at least n* interpretation is available for a numerical NP in predicate position, but not when the same NP is in argument position:

- (12) a. I have four dogs, in fact I have ten.  
 b. ???The inhabitants of the barn are four dogs, and, in fact, there are ten of them.

While (12a) is able to successfully cancel the *exactly n* implication, (12b) is just a correction from *exactly 4* to *exactly 10*. This suggests that the meaning of a cardinal is *exactly n*, but the sentence as a whole has an *at least* entailment, since, while it asserts something about a plurality of *n* entities, it leaves open the possibility that a larger plurality of entities may have the same property. However, since we assume the speaker is as informative as he can be, we do not raise the question whether this group is a part of a group with a bigger cardinality. This explains why (12a) is interpreted with an *at least* entailment, but (12b) is not. Generalizing over contexts of occurrence, we can thus say that cardinals give rise to *exactly n* interpretations when they are embedded under *the* or other determiners or when referred back via an anaphor (contexts which have uniqueness (& maximality) requirement):

$$(13) \quad [[\text{four dogs } \langle_{c,t} \rangle]] = \lambda x. \text{ DOGS}(x) \wedge |x| = 4$$

Otherwise, they receive an *at least n* interpretation:

$$(14) \quad [[\text{four dogs } \langle\langle_{e,t}, t \rangle\rangle]] = \lambda P. \exists x [\text{DOGS}(x) \wedge |x| = 4 \wedge P(x)]$$

If one embraces the pragmatic account of the exact readings of cardinals, the results of the current experiment are expected, given crosslinguistic evidence for derivation of more scalar implicatures with cardinals. However, if one embraces an account relying on *at least* entailment, where interpretations are sensitive to the argument/predicate distinction, or [+/-definite] contexts, then the results are also expected, given that the NPs in all the sentences in the experiment were in argument positions. In other words, the results of the current experiment are compatible with either account, and further research is required to tease these accounts apart.

## 4.2 Discussion

The current experiment shows children assign a precise, upper bound reading to cardinals from early on. However, it does not allow us to say whether the *exactly n* interpretation is a basic interpretation, or whether it is derived via scalar implicatures. While one could easily say that

cardinals are lexically stored with an *exactly n* interpretation, an equally convincing argument would be to say that cardinals are lexical items with an *at least n* interpretation, but that children derive implicatures with them earlier than with quantifiers because they are exposed to a lot of cardinal numbers, and they are often engaged in activities which involve counting (see Papafragou & Musolino 2003), while there is not so much (explicit) engagement with quantifiers. Thus, the strengthening of implicatures with cardinals is favoured by the importance and frequency of counting in child play and education. Another possible explanation for why children do so well with cardinals may be related to the number system of a language. Sarnecka et al. (2007) argue that the conceptual framework of grammatical number supports learning of *one*, *two* and *three*. Sarnecka et al. (2007) make this claim on the basis of corpus analysis (by looking at CHILDES data in English, Russian-which mark singular/plural, and Japanese -which does not), but also on the basis of experiments on children, i.e. Counting and Give-N tasks. Interestingly, the results show English and Russian learners fared better than Japanese learners in the comprehension of number words, irrespective of whether singular/plural cues appeared in the task itself (e.g. *Give two apples* vs. *Give two*). It may be the case that the conceptual system of grammatical number in English, Greek or Romanian boosts *exactly n* performance with cardinals, unlike Japanese or other languages. Such a hypothesis needs further testing. Moreover, it may also be that the conceptual system of grammatical number does not help so much with quantifiers such as *some* versus *all*, as understanding these may require understanding complex relations between sets (e.g. the part-whole relation). In addition, grasping *some* versus *all* requires handling two operators over sets ( $\exists x$  versus  $\forall x$ ), whereas cardinal numbers may be argued to involve no operator at all in their *exactly n* reading (and only the existential operator in their *at least n* reading). In other words, children may produce less implicatures with quantifiers than with cardinals because quantifiers are cognitively more complex.

## 5. Conclusion

To conclude, the current paper presents an experiment on the acquisition of cardinals by Romanian monolingual 5-year-old children, showing that their preferred interpretation is the *exactly n* reading, just as in the case of adults. While one could take this to suggest that children start out with an *exactly n* meaning of cardinals, various arguments can be brought in favour of the view that the exact reading is actually derived via

scalar implicatures. Further research is required to tease apart the two accounts and tip the scales in favour of one account rather than the other.

## Acknowledgements

Thank you to Larisa Avram, Alexandra Cornilescu, Tom Roeper, Anton Benz, Nicole Gotzner and the audience at AICED 21. Thank you to the students from the Faculty of Foreign Languages, University of Bucharest, as well as the educators and children from Waldorf Kindergarten for helping/ taking part in the experiment.

## Funding Information

This material is based upon work supported by the Institute for Research in the Humanities. A research division of The Research Institute of the University of Bucharest (*ICUB*) under Grant TC ICUB-SSU 2804712017 (led by Bleotu Adina Camelia).

## References

- Ariel, Mira. 2006. "A 'just that' lexical meaning for *most*". In *Where Semantics Meets Pragmatics*, edited by Klaus von Heusinger, and Ken Turner, 49-91. Amsterdam: Elsevier.
- Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker, Steve. 2015. "Fitting linear mixed-effects models using lme4". *Journal of Statistical Software* 67 (1): 1-48.
- Cohen, Ariel, and Krifka, Manfred. 2011. "Superlative quantifiers as modifiers of meta-speech acts". In *The Baltic International Yearbook of Cognition, Logic and Communication*, vol. 6, edited by Michael Glanzberg, Barbara Partee, and Jurgis Šķilters, 1-56. Kansas: New Prairie Press.
- Crain, Stephen, and Rosalind Thornton. 1998. *Investigations in Universal Grammar: A Guide to Experiments on Syntax and Semantics*. Cambridge, MA: MIT Press.
- Frege, Gottlob. 1892. "On sense and reference". In *Translations from the philosophical writings of Gottlob Frege*, edited by Peter Geach, and Max Black, 56-78. Oxford: Blackwell.
- Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic.
- Grice, Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

- Horn, Laurence. 1972. "On the Semantic Properties of the Logical Operators in English". PhD diss., UCLA.
- Huang, Yi Ting, Elizabeth Spelke, and Jesse Snedeker. 2013. "What Exactly do Numbers Mean?". *Language Learning and Development* 9 (2): 105-129.
- Hurewitz, Felicia, Anna Papafragou, Lila Gleitman, and Rochel Gelman. 2006. "Asymmetries in the Acquisition of Numbers and Quantifiers". *Language Learning and Development* 2 (2): 77-96.
- Kadmon Nirit. 1987. "On Unique and Non-Unique Reference and Asymmetric Quantification". PhD diss., UMass.
- Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic*. Dordrecht: Kluwer.
- Katsos, Napoleon, and Dorothy Bishop. 2011. "Pragmatic tolerance: implications for the acquisition of informativeness and implicature". *Cognition* 120 (1): 67-81.
- Landman, Fred. 2003. "Predicate-argument mismatches and the Adjectival Theory of Indefinites". In *From NP to DP, Volume 1: The Syntax and Semantics of Noun Phrases*, edited by Martin Coene and Yves D'Hulst, 211-237. Amsterdam: John Benjamins.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, Stephen. 2000. *Presumptive Meanings. The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.
- Noveck, Ira. 2001. "Why children are more logical than adults: Experimental investigations of scalar implicatures". *Cognition* 78 (2): 165-188.
- Papafragou, Anna, and Julien Musolino. 2003. "Scalar implicatures: Experiments at the semantic pragmatics interface". *Cognition* 86 (3): 253-282.
- Pouscoulous, Nausicaa, Ira Noveck, Guy Politzer, and Anne Bastide. 2007. "A developmental investigation of processing costs in implicature production". *Language Acquisition* 14 (4): 347-375.
- R Core Team. 2015. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.
- Rothstein, Susan. 2017. *Semantics for Counting and Measuring*. Cambridge: Cambridge University Press.
- Sarnecka, Barbara W., Valentina G. Kamenskaya, Yuko Yamana, Tamiko Ogura, and Yulia B. Yudobina. 2007. "From Grammatical Number to Exact Numbers: Early Meanings of 'One,' 'Two,' and 'Three' in English, Russian, and Japanese". *Cognitive Psychology* 55 (2): 136-168.

- Stoicescu, Ioana, Anca Sevcenco, and Larisa Avram. 2015. "The acquisition of scalar implicatures in child Romanian". In *Proceedings of Conference on British and American Studies*, edited by Marinela Burada and Oana Tatu, 141-155. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Strawson, Peter Frederick. 1950. "On referring". *Mind* 59: 320-344.

## Appendix

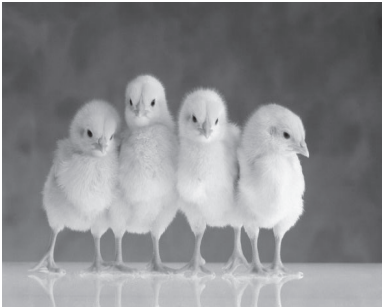
### Test sentences and pictures



*Doi cai au pete albe.*  
‘Two horses have white spots.’



*Doi câini au limba scoasă.*  
‘Two dogs have their tongue out.’



*Trei pui stau în picioare.*  
‘Three chicks are standing up.’



*După gard se află doi măgari.*  
‘Behind the fence there are two donkeys.’



*În pat dorm trei pisici.*  
‘Three kitties are sleeping on the bed.’

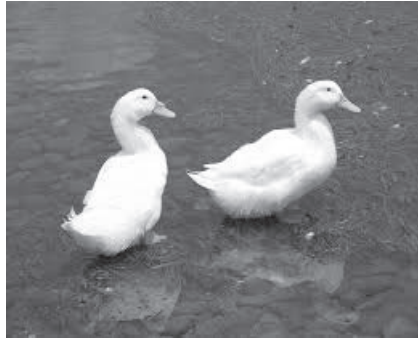


*Pe pajiște trei vaci mănâncă iarbă.*  
‘On the pasture, three cows are eating grass.’

## Control sentences and pictures



*Toate oile sunt albe.*  
'All the sheep are white.'



*Toate rațele se plimbă prin apă.*  
'All the ducks are walking through water.'



*Toti câinii dorm în iarbă.*  
'All the dogs are sleeping in the grass.'



*Unii porcușori sunt roz*  
'Some piglets are pink.'



*Unii pui sunt negri.*  
'Some chicks are black.'



*Unii iepuri țopăie în iarbă.*  
'Some rabbits are hopping in the grass.'



## CHAPTER SIXTEEN

# EXPERIMENTAL DATA ON TELICITY INFERENCES IN ROMANIAN

IOANA STOICESCU

The experimental study in the present paper explores the nature of the completion inferences made by Romanian adult speakers in relation to strict (*make a hat*) and flexible accomplishments (*sew a blouse*) (Wright 2014). It has been claimed that, for strict accomplishments, the idea of completion is an entailment. Faced with the sentence *He made a hat*, interpreters automatically assume that the hat is completed. For flexible accomplishments, completion is not obligatory, being only a pragmatic implicature. The proposition *He sewed a blouse* does not entail that the blouse is finished, the idea of completion being optionally inferred. The experiment used a truth-value judgment task based on written scenarios that presented telic situations that came or did not come to their natural endpoint. The participants were asked to evaluate the truth value of past tense sentences based on strict or flexible accomplishment predicates with indefinite quantized direct objects. The prediction was that there would be a degree of acceptance of incompleteness for flexible accomplishments but that speakers would reject incomplete scenarios with strict accomplishments. The results confirmed this prediction. The speakers' incompleteness acceptance average rates ranged from 36% to 54% for a limited number of predicates, but, for other accomplishments, incompleteness was always rejected. This is evidence that, in Romanian, completion inferences are not obligatory for all telic durative predicates.

Keywords: accomplishments, completion, grammatical aspect, incompleteness, inferences

## 1. Introduction

The traditional analysis of accomplishments defines them as dynamic, durative, telic predicates (*knit a sweater, read a story*) (Vendler 1957, Dowty 1979, Smith 1991/1997). These aspectual properties are derived compositionally at the sentence level, from the joint semantics of the verbs and their arguments (Verkuyl 1993, Rappaport Hovav 2008). The internal argument plays a crucial role in the aspectual interpretation of the sentence. It was argued that quantized internal arguments (definite, indefinite DPs) are associated with telicity (Dowty 1979, MacDonald 2008) because the boundaries of the object affected or created by the event constitute the limits of the event itself (Tenny 1992). In the standard analysis, a sentence with a past tense accomplishment entails that the event has reached its inherent endpoint (1).

(1) I made a cake -> The cake is completed

However, there is agreement in the literature that telic predicates are not a homogeneous class, as it has long been noted that some have “variable telicity”, in the sense that they allow both telic and atelic interpretations. Kennedy and Levin (2008) mention in this category incremental theme predicates (*eat the cake*), degree achievements (*reddden*), directed motion verbs (*ascend*). These authors claim that variable telicity predicates elicit two distinct interpretations because they share scalarity with gradable adjectives, denoting a measure of the “degree to which an object changes relative to a scalar dimension over the course of an event” (Kennedy and Levin 2008, 156). The two interpretations are the following. *He ate a cake* might be read as *He ate a cake completely* or *He ate a significant amount from the cake but did not finish it*. Some researchers have argued that for predicates with variable telicity, completion is not part of propositional meaning but rather a pragmatic inference, an implicature (Wright 2014, Lin 2004). Consequently, it is possible to assign atelic interpretations to such sentences if the context does not strongly suggest the idea of culmination or simply blocks it. Wright (2014) proposes that accomplishments can be divided into two subclasses: strict and flexible accomplishments. For the former, the completion inference is an obligatory entailment, as it is part of the truth conditions of the sentence, while, for the latter, it is only a non-truth-conditional conversational implicature. Strict accomplishments include predicates like *assemble a telescope, convince smb to do smth, create a report, fix the radio, install the program, make a hat, solve the problem*, while flexible accomplishments, supposedly more numerous in

English than the strict variety, include *build a house, drink a cup of coffee, devise a plan, eat a sandwich, establish a business, knit a blanket, organize a trip, paint the barn, peel an orange, read a book, sew a dress, write a story*.

The aim of this pilot study is to investigate whether, in Romanian, the accomplishment class behaves homogeneously with respect to the completion inferences it elicits. The question is whether the division between strict and flexible accomplishments mentioned by Wright (2014) holds in Romanian as well. Are there any accomplishment predicates that allow atelic interpretations? An answer will be provided with the aid of experimental data.

The paper is organized as follows. In the second section, I discuss the analysis of the accomplishment class provided by Wright (2014) and the linguistic criteria and experimental methodology used to differentiate between strict and flexible accomplishments. In the third section, I show the extent to which the diagnostics proposed by Wright for English can be applied to accomplishments in Romanian. In addition, this section includes the description of an experiment which tested Romanian speakers' intuitions related to accomplishment sentences in complete and incomplete scenarios. The section ends with a presentation of the results and their discussion. The last section concludes the paper.

## **2. Strict vs. flexible accomplishments in English - Wright (2014)**

Wright defines strict accomplishments as durative telic predicates which are “unambiguously endpoint-inclusive”, while flexible accomplishments “permit either an endpoint-inclusive or non-endpoint-inclusive reading” (2014, 35). This division is based on linguistic diagnostics and experimental evidence. The tests that differentiate the two subclasses are summarized in Table 1 (adapted from Table 8 in Wright 2014, 162) and elaborated on in Sections 2.1.1 to 2.1.5. Atomic minimal-eventivity will be tackled in Section 3.4.3.

	FAs	SAs
<i>For x time</i> expressions	✓	X
<i>In x time</i> expression	✓	✓
While-continuations relax telicity	✓	X
Pauses and continuations	✓	X
Can be atomic minimal-eventive	✓	X

**Table 1 Diagnostics for flexible (FAs) vs strict accomplishments (SAs)**

## 2.1 For x time/in x time

Flexible accomplishments may felicitously occur with both *in x time* and *for x time* adverbials. In the context of the completive *in x time* adverbial, they have a telic interpretation and suggest completion, and, distributed with the durative *for x time* adverbial, they have an atelic interpretation, and are associated with the idea of a process that does not reach its inherent endpoint (2).

- (2) a. Jay ate a sandwich in a few minutes. (examples 66-67 in Wright 2014, 46)
- b. Dylan knitted a sweater in a few days.
- c. Jay ate a sandwich for a few minutes, but he realized he didn't like it.
- d. Dylan knitted a sweater for a few hours, but gave up in frustration.

Strict accomplishments only appear in the context of *in x time* adverbials (3a-b), and typically do not co-occur with *for x time* adverbials (3c-d).

- (3) a. Karla repaired the watch in half an hour. (examples 68-69 in Wright 2014, 46)
- b. Samara made a box in a couple of minutes.
- c. #Karla repaired the watch for half an hour, but gave up in frustration.
- d. #Samara made a box for a couple of minutes, but soon grew bored and stopped.

## 2.2 While-continuations

Flexible accomplishments elicit a completive interpretation in (4a-c) but do not require it in sentences including a *while* subordinate with a past tense or an activity predicate in the participle (4d-f). The *while*-subordinate forces an open reading of the accomplishment and the sentences in (4d-f) are coerced activities - the endpoint is not necessarily visible. As completion is not required with the *while*-clause, the example in (4d) could be continued as in (4g), where the endpoint is cancelled (Wright 2014, 47).

- (4)
- a. Jay played a sonata. (examples 70-71 in Wright 2014, 46-47)
  - b. Lynn read *War and Peace*.
  - c. Chris sewed a jacket.
  - d. Jay played a sonata while the maître d' stole his tips.
  - e. Lynn read a novel while Jack napped on the sofa.
  - f. Chris sewed a jacket while waiting for the bus.
  - g. Jay played a sonata while the maître d' stole his tips. Jay stopped and confronted him.

With strict accomplishments, the interpretation is closed – the event reaches its final point before the end of the interval denoted by the *while*-clause (5).

- (5)
- a. Karla repaired the watch while watching the news.
  - b. Samara made a box while Dusty took pictures for the newsletter.
  - c. Leon solved the problem while others were panicking. (example (73) in Wright 2014, 47)

## 2.3 “Pauses and Continuations of Action”

Wright (2014) argues that flexible accomplishments allow interruptions or pauses, followed by the resumption of the event (6):

- (6)
- a. Jay played a sonata, paused to take a sip of water, and resumed playing it.
  - b. Lynn read *War and Peace*, took some notes, then read some more.
  - c. Chris sewed a jacket, stopped to check the pot roast, then sewed a little more. (example 74 in Wright 2014, 47)

Strict accomplishments do not allow interruptions, nor the possibility for the event to resume (7)

- (7) a. #Karla repaired the watch, took a lunch break, then repaired it some more.  
 b. #Samara made a box, paused to read the instructions, then resumed making it.  
 c. #Leon solved the problem, got called away, then came back and continued solving it. (example 75 in Wright 2014, 48)

## 2.4 Experimental evidence for the distinction between strict and flexible accomplishments

Wright (2014) conducted two experiments in which strict and flexible accomplishments were evaluated in scenarios in which telic situations came to their inherent endpoints or not. Both experiments included truth-value judgment tasks but two experimental methods were used – either presentation of short movies or written scenarios. In both experiments English speakers accepted sentences with flexible accomplishments like *Ray read a book* in non-completed scenarios but rejected strict accomplishments like *Ray assembled a telescope* in such scenarios. In Experiment 2, for the flexible accomplishments *peel an orange*, *write a story*, *paint a barn*, *eat a sandwich*, *drink a cup of coffee*, the acceptance rate in non-completed scenarios exceeded 50%.

## 2.5 Completion is a cancellable implicature with flexible accomplishments

Wright (2014) accounts for the high acceptance rates of flexible accomplishments in incompletion contexts by arguing that the completion inferences made with this class are conversational implicatures, not entailments.<sup>1</sup> As such they can be cancelled in the right context (8b, d) but also strengthened (9) (Coppock 2012 *apud* Wright 2014, 120).

- (8) a. Ulbricht wrote a memorandum. (examples (15)-(16) in Wright 2014, 117).  
 b. Ulbricht wrote a memorandum until his typewriter jammed.  
 c. Chris sewed a dress.  
 d. Chris sewed a dress, but ran out of material before she could finish.

- (9) a. Ray ate a sandwich, and he ate all of it. (examples (31)-(34) in Wright 2014, 120).  
 b. Simon sewed a dress. In fact, he finished sewing it.  
 c. I drank the cup of coffee, and I drank it all.  
 d. Noelle read *War and Peace*; furthermore, she read it all.

The completion conversational implicature is based on the Gricean maxim of quantity (Grice 1975). When addressees hear sentences like (8a, c), they expect speakers to have been as informative as possible about the trajectory of the event and the extent of its progress towards its natural limit. Since the relevant sentences include quantized internal arguments, which are naturally assumed to refer to whole objects, not incomplete ones, then hearers infer that the endpoint of the event has been reached.

### 3. Strict vs. flexible accomplishments in Romanian

In this section, I will present the case for the relevance of the distinction between strict and flexible accomplishments in Romanian. I will list and discuss the diagnostics presented by Wright in order to determine whether their use can be extended to Romanian. Then I will present the results of a replica of Wright's Experiment 2, which confirms that the division is tenable in Romanian as well. The judgments used for the discussion of the diagnostics are based on informal questionnaires distributed to ten native speakers of Romanian. As this is a pilot exploratory study, I started off by using the translations of the English predicates used by Wright, although I am aware that the lexicalized aspectual properties of the Romanian predicates might be different.

#### 3.1 For *x* time

Romanian verb constellations based on *a mânca* "eat", *a bea* "drink", *a curăța* "peel", *a spăla* "wash" behave like English flexible accomplishments, co-occurring with *for x time*. The sentences in (10) with the durative adverbial were accepted by the informants<sup>2</sup>.

- (10) a. Ion a mâncat un sandvici timp de câteva minute.  
 Ion has eaten a sandwich for several minutes  
 "Ion ate a sandwich for a few minutes."  
 b. Ion a băut o bere timp de câteva minute.  
 Ion has drunk a beer for several minutes.  
 "Ion drank a beer for a few minutes."

- c. Ion a curățat un măr timp de un minut.  
 Ion has peeled an apple for a minute  
 “Ion peeled an apple for a minute.”
- d. Mihai a spălat o farfurie timp de două minute.  
 Mihai has washed a plate for two minutes.  
 “Mihai washed a plate for two minutes.”

However, the *for x time* adverbial was deemed less felicitous in interruption contexts, although it was not rejected absolutely (11).

- (11) a. ?Ion a mâncat un sandwich timp de câteva minute,  
 Ion has eaten a sandwich for several minutes  
 dar apoi l-a lăsat pe masă.  
 but then CL.ACC.3.SG.NEUT. has left on table  
 “Ion ate a sandwich for a few minutes, but then left it on the table.”
- b. ?Ion a desenat un copil timp de cinci minute,  
 Ion has drawn a child for five minutes,  
 dar i s-a rupt creionul.  
 But CL.DAT.3.SG.MASC. REFL. has broken pencil-the  
 “Ion drew a child for five minutes but then his pencil broke.”

Sentences with the Romanian equivalents of Wright’s strict accomplishments *assemble*, *compose*, *install* and *for x time* were considered acceptable, although not perfect, in informal questionnaires (12). About half of the informants found these examples worse than their flexible accomplishment counterparts in (10), but on average the two types of examples were treated the same way.<sup>3</sup>

- (12) a. Mihai a asamblat un dulap timp de o zi.  
 Mihai has assembled a cupboard for a day.  
 “Mihai assembled a cupboard for a day”
- b. Andrei a compus un cântec timp de o săptămână.  
 Andrei has composed a song for a week.  
 “Andrei composed a song for a week.”
- c. Gabi a instalat un program timp de jumătate de oră.  
 Gabi has installed a program for half of hour  
 “Gabi installed a program for half an hour.”



With *for x time* and interruptions, the strict accomplishments examples are further degraded (13). The informants found such examples less acceptable than the flexible accomplishments in (11).

- (13) a. ??Andrei a compus un cântec timp de  
 Andrei has composed a song for  
 o săptămână, dar nu l-a terminat.  
 a week but not CL.ACC.3.SG.M-has finished  
 “Andrei composed a song for a week, but he did not finish it.”
- b. ??Mihai a asamblat un dulap timp de o zi,  
 Mihai has assembled a cupboard for a day  
 dar nu l-a terminat.  
 but not CL.ACC.3.SG.M. has finished  
 “Mihai assembled a cupboard for a day but he did not finish it.”
- c. ??Gabi a instalat un program timp de jumătate de oră,  
 Gabi has installed a program for half of hour  
 dar nu a terminat.  
 but not has finished  
 “Gabi installed a program for half an hour but did not finish it.”

### 3.2 *While*-continuations

Wright argues that flexible accomplishments lose their completive interpretation in sentences with a *while* subordinate, and thus allow interruptions (4g), unlike strict accomplishments, which always suggest completion in this context and would prohibit interruptions. In Romanian, however, flexible accomplishments were not more permissive of interruptions than their strict counterparts. Both classes of predicates were given a similar assessment, as being slightly unacceptable.

- (14) a. Maria a desenat un copil în timp ce sora ei  
 Maria has drawn a child in time what sister.DEF her  
 asculta muzică. ?Maria s-a oprit  
 listen.IMPERF.3SG. music. Maria REFL-has stopped  
 din desenat și i-a spus  
 from draw.SUP and CL.DAT.3.G.FEM-has said

că muzica e prea tare.  
that music.DEF is too loud

“Maria drew a child while her sister was listening to music. Maria stopped drawing and told her that the music was too loud.”

- b. Sarah a scris o scrisoare în timp ce  
Sarah has written a letter in time what  
Silvana dormea. ?Sarah s-a oprit  
Silvana sleep.IMPERF.3SG. Sarah REFL-has stopped  
din scris și a trezit-o.  
from write.SUP and has woken-CL.ACC.3.SG.FEM.  
“Sarah wrote a letter while Silvana was sleeping. Sarah stopped writing and woke her up.”

- (15) a. Ion a reparat un robinet în timp ce sora lui  
Ion has fixed a tap in time what sister.DEF his  
asculta știrile. ?Ion s-a oprit din  
listen.IMPERF.3SG. news.DEF. Ion REFL-has stopped from  
reparat și i-a spus să stingă radioul.  
fix.SUP and CL.DAT.3.SG.FEM-has said SUBJ turn-off radio.DEF  
“Ion fixed a tap while his sister was listening to the news. Ion stopped writing and told her to switch off the radio.”
- b. Ion a instalat un program în timp ce sora lui  
Ion has installed a program in time what sister.DEF his  
asculta știrile. ??Ion s-a oprit  
listen.IMPERF.3SG. news.DEF. Ion REFL-has stopped  
din instalat și i-a spus  
from install.SUP and CL.DAT.3.SG.FEM-has said  
să stingă radioul.  
SUBJ turn-off radio.DEF  
“Ion installed a program while his sister was listening to the news. Ion stopped installing and told her to switch off the radio.”

The *while*-continuation test is not relevant for Romanian because the *perfect compus* in the main clause induces a perfective reading, regardless of the type of accomplishment used. The most natural reading for the examples above is that the main event reached completion before the end of the secondary event described in the *while*-clause. That is why a continuation which suggests that the main event did not reach its natural endpoint is odd for both types of accomplishments.

### 3.3 “Pauses and Continuations of Action”

Wright claims that flexible accomplishments allow interruptions or pauses, followed by the resumption of the event, while strict accomplishments do not (6-7). In Romanian, both types of accomplishments were assessed as being unacceptable but flexible accomplishments (16) fared slightly better than strict accomplishments (17).

- (16) ?Ion a mâncat un sandvici, s-a oprit să bea apă și apoi l-a mâncat în continuare.  
 “Ion ate a sandwich, stopped to drink some water and then continued eating it.”
- (17) ??Silvia a făcut un om de zăpadă, s-a oprit să se uite la desene animate și apoi l-a făcut în continuare.  
 “Silvia made a snowman, she stopped to watch some cartoons then continued making it.”

To sum up, the diagnostics listed by Wright did not yield as conclusive results for Romanian accomplishments as they did in English. I used informal grading questionnaires while Wright reports his own judgments, so the difference in methodology is likely to have played a part. Both strict and flexible accomplishments were found acceptable in sentences with *for x time* by Romanian speakers. Only some speakers evaluated strict accomplishments as less acceptable in such sentences. Strict accomplishments were considered less acceptable than flexible accomplishments when tested with the *for x time* and interruption diagnostic. The *while*-continuation sentences did not highlight any differences between the two categories of predicates because of the unambiguously perfective reading of the *perfect compus* in this context. The pauses and continuations test evinced a distinction between strict and flexible accomplishments in that strict accomplishments were found to be less acceptable. The findings are summarised below in Table 2. The intuitions of Romanian speakers also highlight the existence of a distinction between the English past tense and the Romanian *perfect compus*. The latter generates more rigidly perfective readings with telic predicates than the former. This might be due to the fact that the Romanian *perfect compus* is aspectually sensitive, in the sense that it requires a bounded eventuality as its complement (Crăniceanu 2002), while the English past tense is aspectually transparent and does not impose such restrictions (De Swart 1998). I have to leave the articulation of this intuition for future research.

	FAs	SAs
<i>For x time</i>	✓	✓ (speaker variability)
<i>For x time + interruption</i>	?	??
While-continuations	?	?
Pauses and continuations	?	??

**Table 2 Diagnostics for flexible (FAs) vs strict accomplishments (SAs) in Romanian**

### 3.4 Experimental data

The hypothesis that flexible accomplishments constitute a real class was tested by Wright experimentally using a truth value judgment task (Crain and Thornton 1998). He presented English speakers with scenarios in which telic situations described with either flexible or strict accomplishments came to completion or did not reach their inherent endpoint. The speakers then assessed the truth value of sentences with the respective past tense strict/flexible accomplishments. Wright found that English speakers accepted flexible accomplishments and rejected strict accomplishments in the incomplete scenarios. The rate of acceptance for flexible accomplishments was very high, which constituted evidence that the completion inference is not obligatory with these predicates. In this section, I will present a replica of Wright's Experiment 2 that uses the same methodology, and some of the predicates used by Wright in scenarios adapted for Romanian. The question is whether Romanian speakers accept flexible accomplishments in incomplete scenarios. If they do, this would confirm Wright's account, namely that completion is not compulsory, hence not truth-conditional for certain durative telic predicates, and that speakers have to have pragmatic reasons to think about completion.

#### 3.4.1 Method, materials and participants

The present experiment used a truth value judgment task. A written questionnaire, comprising 21 scenarios in which the events either reached their natural endpoint or stopped short of it, was presented. After the participants read the scenario, they evaluated the truth value of three sentences that were related to the story, one of which was the test item (20), while the others were fillers (one true, one false). The fillers were necessary

in order to prevent the respondents from realizing that the experiment tested their reaction to the completion-incompletion contrast. Had they figured out the purpose of the task, they would have been inclined to reject all the test items in the incompletion scenarios and accept them in the completion scenarios. Following the methodology in Wright (2014), I double-checked whether the participants realized what the focus of the experiment was by asking them to write a brief commentary about the experiment and what they thought it had been about. Most participants thought it was a test of their capacity to pay attention to details, which proved that the filler items had been effective in concealing the point of the experiment (although there were some participants who inferred its real goal). There are examples of the completion and incompletion scenarios in (18-19) and the test sentence is illustrated in (20). Each completion scenario had an incompletion counterpart built around the same predicate. Two test item lists were compiled so that no list would include both completion and incompletion scenarios for the same test item – each list was administered to half of the participants. Both lists were randomized. The order in which the two fillers and the test item were presented was also randomized.

(18) *Completion scenario*

Mihaelei îi era foame. Din fericire, avea la ea un sandvici cu pâine cu măsline și brânză. L-a scos din geantă. Sandviciul arăta minunat. A mușcat cu poftă din el de multe ori. Îi plăcea atât de mult încât nu se putea opri. A înghițit ultimele firimituri și apoi s-a dus să se spele pe mâini.

*Mihaela was hungry. Fortunately, she had a cheese and olive bread sandwich on her. She took it out of her bag. It looked great. She took several bites. It was so good she couldn't stop. She ate the last crumbs and then went to wash her hands.*

(19) *Incompletion scenario*

Mihaelei îi era foame. Din fericire, avea la ea un sandvici cu pâine cu măsline. L-a scos din geantă. Sandviciul arăta minunat. A mușcat cu poftă din el de multe ori, apoi a simțit o durere puternică de măsea. După câteva secunde, și-a dat seama că în sandvici fusese un sambure tare de măslină. Suparată, a aruncat sandviciul la gunoi.

*Mihaela was hungry. Fortunately, she had a cheese and olive bread sandwich on her. She took it out of her bag. It looked great. She took several*

*bites, but then she got a strong toothache. After a few seconds, she realized she had bit on a hard olive pit. She was so angry she threw away the sandwich in the garbage.*

(20) *Test sentence:*

Mihaela a mâncat un sandvici.  
 “Mihaela ate a sandwich.”

The predicates tested were the following (based on Wright 2014): a) 11 (hypothesized) **flexible accomplishments**: *a curăța un măr* “peel an apple”, *a bea o bere* “drink a beer”, *a vopsi un gard* “paint a fence”, *a spăla un câine* “wash a dog”, *a citi o poveste* “read a story”, *a coase o ie* “sew a blouse”, *a scrie o scrisoare* “write a letter”, *a desena un copil* “draw a child”, *a picta un portret* “paint a portrait”, *a mânca un sandvici* “eat a sandwich”, *a șterge o masă* “wipe a table”; b) 6 (hypothesized) **strict accomplishments**: *a asambla un dulap* “assemble a cupboard”, *a face o ciorbă* “make a soup”, *a compune un cântec* “compose a song”, *a instala un program* “install a program”, *a rezolva o problemă* “solve a problem”, *a repara un frigider* “fix a fridge”. The list of test items included predicates of creation, affected object predicates and predicates of consumption.

The list of items also included 4 telic predicates which served as control items (expected not to allow non-completive readings, given the relevant findings in the literature (Van Hout et al. 2017): *a rupe o creangă* “break a tree branch”, *a deschide o fereastră* “open a window”, *a închide o ușă* “close a door”, *a stinge o lumânare* “blow out a candle”. These predicates are affected object predicates and usually have non-durative interpretations (which qualifies them as achievements in the standard view – Smith 1997).

There were 4 experimental conditions: a) completion scenario – flexible accomplishment, b) incompleteness scenario – flexible accomplishment; c) completion scenario - strict accomplishment; d) incompleteness scenario - strict accomplishment, and 1 control condition.

The participants were 20 adult speakers of Romanian (students of the University of Bucharest). The first version of the list was administered to 11 speakers and the second version of the list was administered to 9 speakers.

### 3.4.2 Results

The results of the questionnaire are given in Table 3, which presents the acceptance rates for each test and control item in the completion and incompleteness scenarios.

<i>FA acceptance rates (%)</i>		
	<b>Completion</b>	<b>Incompletion</b>
Peel an apple	100	0
Drink a beer	100	0
Paint a fence	90	11
Wash a dog	100	22
Read a story	90	0
Sew a blouse	89	9
<b>Write a letter</b>	100	<b>55</b>
<b>Draw a child</b>	89	<b>55</b>
<b>Paint a portrait</b>	78	<b>36</b>
<b>Eat a sandwich</b>	100	<b>36</b>
<b>Wipe a table</b>	89	<b>46</b>
<i>SA acceptance rates (%)</i>		
Solve a problem	100	9
Assemble a cupboard	100	0
Make a soup	100	18
Fix a fridge	55	11
Compose a song	91	0
Install a program	91	0
<i>Control items acceptance rates (%)</i>		
Blow out a candle	89	0
Break a tree branch	100	0
Open a window	100	0
Close a door	78	45

**Table 3** Acceptance rates

As can be seen, some hypothesized flexible accomplishment predicates (*a curăța un măr* “peel an apple”, *a bea o bere* “drink a beer”, *a vopsi un gard* “paint a fence”, *a spăla un câine* “wash a dog”, *a citi o poveste* “read a story”, *a coase o ie* “sew a blouse”,) were generally not accepted in the incompleteness scenarios as the acceptance rates are low – at most 20% (this means there were only 1 or 2 acceptance responses for these predicates).

For five predicates (*a scrie o scrisoare* “write a letter”, *a desena un copil* “draw a child”, *a picta un portret* “paint a portrait”, *a mânca un sandvici* “eat a sandwich”, *a șterge o masă* “wipe a table”) the acceptance rates in incompleteness scenario are higher, ranging between 35% - 55%, which indicates a much higher number of responses in which the test items were judged true even if the event had not reached its endpoint. These predicates would fall in the flexible accomplishment class in Romanian, unlike their counterparts mentioned before.

The predicates that Wright regarded as strict accomplishments were consistently rejected in incompleteness scenarios, at levels comparable to the control items. There is one exception (the hypothesized SA *a repara un frigider* “fix a fridge” paradoxically has a low acceptance rate in the completion scenario). The control item *a închide o ușă* “close a door” was frequently rejected in the incompleteness situation. The unexpected results for these two predicates are likely due to imperfections in the scenarios which made the participants overthink.

### 3.4.3 Discussion

The results confirm that, in Romanian, there are predicates that behave like Wright’s flexible accomplishments in that speakers may accept past tense sentences with such predicates in incompleteness scenarios. I have identified five such predicates for now, but the experiment needs to be extended to other predicates and it should be administered to a larger number of speakers. This is evidence in the support of the hypothesis that the accomplishment class is not homogeneous with respect to completion inferences – for some predicates, speakers may assign both culminating and non-culminating interpretations, while for others the completion inference is always made and is the only option. Following Wright (2014), it could be argued that completion is not an entailment with the flexible accomplishment predicates rather a cancellable implicature. However, in order to demonstrate this with a reasonable degree of certainty it would be necessary to identify some contextual factors which would trigger the implicature – this will be the focus of future research.

Wright’s account of the English data is based on the notion of **atomic minimal eventivity**. The ambiguity of flexible accomplishments is linked to the fact that they may contain **atomic minimal events** (along the lines of Rothstein 2008), while strict accomplishments do not. According to Wright (2014, 107), “atomic minimal events are iterated discrete subevents of complex events such that each atomic minimal event is a token of the same type.” For instance, the minimal atomic event related to walking is the step



(Wright 2014), for jumping, one jump (Rothstein 2008). Events that comprise atomic minimal events can be intuitively divided into minimal event units that have a beginning and an end (Rothstein 2008, 12) – while *jump* does comprise of atomic minimal events, *run* does not.

Wright distinguishes between flexible accomplishments that comprise of atomic minimal events, and those that do not. For instance, *sew a blouse* falls in the first category, as it involves repeatedly pushing and pulling thread through a piece of cloth, as does *drink a cup of tea*, which amounts to taking consecutive sips. For this sort of FA, once a full atomic minimal event has been perceived, the speaker can safely assume that the predicate applies. If one sip has been taken, drinking a cup of tea has already occurred (Wright 2014, 163). However, Wright acknowledges the fact that there are some FAs that do not consist of atomic minimal predicates because they designate situations with a heterogeneous event structure (*organize a meeting, build a garage*) (Wright 2014, 131). Completion inferences are influenced by the salience of the perceived result of the event in this case. Even if the situation has not reached its final endpoint, if the effect of the event on the internal argument is sufficiently salient for the speaker, then the past accomplishment sentence might be accepted. Wright suggests that some speakers may actually be making an acceptability judgment rather than a truth value judgment and, in fact, some speakers may be more sensitive to truth/falsity, others to the visibility of the effect of the event.

Coming back to the Romanian data, the list of the alleged FAs tested included both atomic minimal eventive predicates (*a curăța un măr* “peel an apple”, *a mânca un sandvici* “eat a sandwich”), and also heterogeneous ones (*a citi o poveste* “read a story”). Atomic minimal eventhood did not necessarily lead to greater number of responses in which the test items received a non-completive interpretation - *a curăța un măr* “peel an apple”, *a bea o bere* “drink a beer”, *a coase o ie* “sew a blouse” were not accepted in the incompleteness scenarios, while *a scrie o scrisoare* “write a letter”, *a mânca un sandvici* “eat a sandwich”, *a șterge o masă* “wipe a table” were allowed. Both types of predicates are minimally eventive, so minimal atomic eventivity cannot be the only factor that explains the Romanian speakers’ judgments. As for the second factor proposed by Wright, the salience of the effect on the object, it is not possible to evaluate its influence in this study because the experiment did not have a built-in way to gauge it.

## 4. Conclusions

This study has investigated the hypothesis—previously put forth for English—that accomplishments are subdivided into two classes with distinct

properties, flexible and strict accomplishments. It has highlighted the fact that the diagnostics proposed by Wright (2014) have some applicability in Romanian, limited by the differences between the English past tense and the Romanian *perfect compus*. The diagnostics evinced distinct degrees of acceptability for the two subclasses of predicates, in contexts that involve the interruption of the event. This result was corroborated by experimental evidence that Romanian speakers accept past tense sentences that include (non exclusively) minimal atomic event predicates in incompletion scenarios. Further research should extend the empirical coverage of the experiment and diagnostics and should find the means to test the idea that the nature of the completion inference is pragmatic rather than semantic in the case of flexible accomplishments.

## References

- Coppock, Elizabeth. 2012. Lecture 9: Conventional Implicatures. Unpublished lecture notes. Retrieved March 18, 2014: <http://eecoppock.info/PragmaticsSoSe2012/lecture9-potts-2up.pdf>.
- Crain, Stephen, and Rosalind Thornton. 1998. *Investigations in Universal Grammar: A Guide to Experiments in the Acquisition of Syntax and Semantics*. Cambridge, MA: MIT Press.
- Crăniceanu, Ilinca. 2002. "Aspect and Coercion in Romanian Perfect Compus and Imperfect." *Balkanistica* 15: 9-146.
- Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.
- Grice, H. Paul. 1975. "Logic and Conversation". In *Speech Acts*, edited by Peter Cole, and Jerry L. Morgan, 41–58. New York: Academic Press.
- Hacohen, Aviya. 2009. "On the Acquisition of (Hebrew) Compositional Telicity". PhD diss., Ben-Gurion University of the Negev.
- Kearns, Kate. 2000. *Semantics*. New York: Palgrave MacMillan.
- Kennedy, Christopher, and Beth Levin. 2008. "Measure of Change: The Adjectival Core of Degree Achievements". In *Adjectives and Adverbs: Syntax, Semantics and Discourse*, edited by Louise McNally and Christopher Kennedy, 156-182. Oxford: Oxford University Press.
- Lin, Jimmy. 2004. *Event Structure and The Encoding of Arguments: The Syntax of the Mandarin and English Verb Phrase*. PhD diss., Massachusetts Institute of Technology.
- MacDonald, Jonathan. 2008. *The syntactic Nature of Inner Aspect*. Amsterdam/Philadelphia: John Benjamins.
- Rappaport-Hovav, Malka. 2008. "Lexicalized Meaning and the Internal Temporal Structure of Events". In *Theoretical and Crosslinguistic*

- Approaches to the Semantics of Aspect*, edited by Susan Rothstein, 13-42. John Benjamins.
- Rothstein, Susan. 2008. "Two Puzzles for a Theory of Lexical Aspect: Semelfactives and Degree Achievements". In *Language, Context and Cognition: Event Structures in Linguistic Form and Interpretation*, edited by Johannes Dölling, Tatjana Heyde-Zybatow, and Martin Schäfer, 175-197. Berlin: Walter de Gruyter.
- Smith, Carlota. 1997. *The Parameter of Aspect*. Dordrecht: Kluwer.
- de Swart, Henriette. 1998. "Aspect Shift and Coercion". *Natural Language and Linguistic Theory* 16 (2): 347-385.
- Tenny, Carol. "The Aspectual Interface Hypothesis". In *Lexical Matters*, edited by Ivan A. Sag and Anna Szabolcsi, 1-28. Centre for the Study of Language and Information.
- Van Hout, Angeliek, María Arche, Hamida Demirdache, Isabel García del Real, Ainara García Sanz, Anna Gavarró, Lucía Gómez Marzo, Saar Hommes, Nina Kazanina, Jinhong Liu, Oana Lungu, Fabienne Martin and Iris M. Strangmann. 2017. "Agent Control and the Acquisition of Event Culmination in Basque, Dutch, English, Spanish and Mandarin". In *Proceedings of the 41st Annual Boston University Conference on Language Development*, edited by Maria LaMendola and Jennifer Scott, 323-332. Somerville, MA: Cascadilla Press.
- Vendler, Zenon. 1957. "Verbs and times". *The Philosophical Review* 66: 143-160.
- Verkuyl, Henk. 1993. *A Theory of Aspectuality*. Cambridge: Cambridge University Press.
- Wright, Tony. 2014. "Strict vs. Flexible Accomplishment Predicates". PhD diss., University of Texas at Austin.

## Notes

<sup>1</sup> Entailments are deductions that are inherent in the propositional content of a sentence. *He made a box* entails *He completed the box*. Conversational implicatures are inferences that are not triggered by the semantic content of a sentence, being the result of the application of pragmatic norms such as the Gricean maxims and the Principle of Cooperation (Grice 1975, Kearns 2000).

<sup>2</sup> An anonymous reviewer raises the issue that predicates with direct object DPs with indefinite articles (10) are not really compatible with durative time adverbials in Romanian because they suggest completion. The reviewer also asks whether there are any interpretative differences between sentences with definite vs indefinite article direct object DPs. Clearly, there is a lot of speaker variability when it comes to the felicitousness of sentences with durative adverbials such as those in (10). Our questionnaire informants found these examples, on average, acceptable. Additionally,

---

the results of the experiment itself show that non-completed readings are allowed for sentences such as (20). If non-completed readings are possible for these predicates in sentences without the adverbials, it is not unreasonable to assume that these interpretations are accessed for sentences with durative adverbials as well. However, our study used only a limited number of informants, and further surveys should be conducted with a higher number of participants. Wright (2014) claimed that sentences such as *John ate a cake for ten minutes* are acceptable in English. With respect to the impact of the definite vs indefinite article within the DP object on the interpretation of sentences such as (10), the issue has not been studied for Romanian yet. For Hebrew, Hacoen (2009) found that the rates for the acceptance of predicates with definite vs. indefinite direct objects for incomplete situations were similar.

<sup>3</sup> Some Romanian speakers rejected (12b), claiming that the meaning of the verb *compune* 'compose' focuses the final stages of a creative endeavour.

## CONTRIBUTORS

Andrei A. Avram, Bucharest University, [andrei.avram@lls.unibuc.ro](mailto:andrei.avram@lls.unibuc.ro)

Gabriela Bîlbîie, Bucharest University, [gabriela.bilbiie@lls.unibuc.ro](mailto:gabriela.bilbiie@lls.unibuc.ro)

Adina-Camelia Bleotu, Bucharest University, [adina.bleotu@lls.unibuc.ro](mailto:adina.bleotu@lls.unibuc.ro)

Ana Regina Calindro, Federal University of Rio de Janeiro,  
[anacalindro@letras.ufjf.br](mailto:anacalindro@letras.ufjf.br)

Elena-Raluca Constantin, Military Technical Academy of Bucharest,  
[raluconstantin@yahoo.co.uk](mailto:raluconstantin@yahoo.co.uk)

Ionuț Geană, Romanian Academy, “Iorgu Iordan—Al. Rosetti” Institute of  
Linguistics; Bucharest University; School of International Letters and  
Cultures, Arizona State University; [ionut.geana@litere.unibuc.ro](mailto:ionut.geana@litere.unibuc.ro)

Ion Giurgea, Romanian Academy, “Iorgu Iordan—Al. Rosetti” Institute of  
Linguistics, [iongiurgea@yahoo.com](mailto:iongiurgea@yahoo.com)

Jutta Hartmann, Bielefeld University, [jutta-maria.hartmann@uni-bielefeld.de](mailto:jutta-maria.hartmann@uni-bielefeld.de)

Dana Isac, Concordia University, [dana.isac@concordia.ca](mailto:dana.isac@concordia.ca)

Mara Panaitescu, Bucharest University, [mara.panaitescu@lls.unibuc.ro](mailto:mara.panaitescu@lls.unibuc.ro)

Ioana Stoicescu, Bucharest University, [ioana.stoicescu@lls.unibuc.ro](mailto:ioana.stoicescu@lls.unibuc.ro)

Balázs Surányi, Pázmány Péter Catholic University; ELKH – Hungarian  
Research Centre for Linguistics, [suranyi.balazs@nytud.hu](mailto:suranyi.balazs@nytud.hu)

Mihaela Tănase-Dogaru, Bucharest University,  
[mihaela.dogaru@lls.unibuc.ro](mailto:mihaela.dogaru@lls.unibuc.ro)

Gergő Turi, Pázmány Péter Catholic University; ELKH – Hungarian  
Research Centre for Linguistics, [turi.gergo@nytud.hu](mailto:turi.gergo@nytud.hu)

Alina Tigău, Bucharest University, [alina.tigau@lls.unibuc.ro](mailto:alina.tigau@lls.unibuc.ro)

Veronica Tomescu, Bucharest University, [veronica.tomescu@lls.unibuc.ro](mailto:veronica.tomescu@lls.unibuc.ro)

Mihaela Zamfirescu, Bucharest University,  
[mihaela.zamfirescu@lls.unibuc.ro](mailto:mihaela.zamfirescu@lls.unibuc.ro)