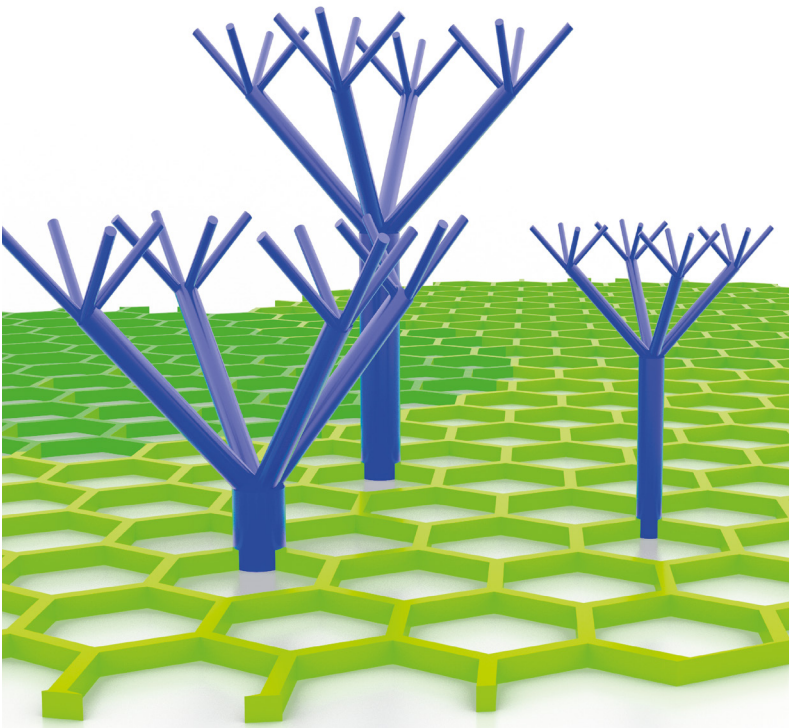
Edited by
Markéta Janebová, Joseph Emonds, and Ludmila Veselovská

Language Use and Linguistic Structure

Proceedings of the Olomouc Linguistics Colloquium 2021



OLOMOUC MODERN LANGUAGE SERIES VOL. 9

Language Use and Linguistic Structure

Proceedings of the Olomouc Linguistics Colloquium 2021

Edited by Markéta Janebová, Joseph Emonds, and Ludmila Veselovská

Palacký University
Olomouc

2022

OLOMOUC MODERN LANGUAGE SERIES (OMLS) publishes peer-reviewed proceedings from selected conferences on linguistics, literature, and translation studies held at Palacký University, Olomouc, Czech Republic.

Published so far:
OMLS, Vol. 1: Teaching Translation and Interpreting in the 21st Century (2012)
OMLS, Vol. 2: Tradition and Trends in Trans-Language Communication (2013)
OMLS, Vol. 3: Language Use and Linguistic Structure. Proceedings of the Olomouc Linguistics Colloquium 2013 (2014)
OMLS, Vol. 4: Complex Visibles Out There. Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure (2014)
OMLS, Vol. 5: Interchange between Languages and Cultures: The Quest for Quality (2016)
OMLS, Vol. 6: Language Use and Linguistic Structure. Proceedings of the Olomouc Linguistics Colloquium 2016 (2017)
OMLS, Vol. 7: Language Use and Linguistic Structure. Proceedings of the Olomouc Linguistics Colloquium 2018 (2019)
OMLS, Vol. 8: Teaching Translation vs. Training Translators. Proceedings of the Translation and Interpreting Forum Olomouc (2022)

OLOMOUC MODERN LANGUAGE SERIES
Vol. 9

Language Use and Linguistic Structure

Proceedings of the Olomouc Linguistics Colloquium 2021

organized by

Department of English and American Studies
Faculty of Arts, Palacký University, Olomouc, Czech Republic
June 10–12, 2022

Edited by Markéta Janebová, Joseph Emonds, and Ludmila Veselovská

Palacký University
Olomouc
2022

5HYLHZHU RI WKH YROXPH 0RMPtU 'RþHNDO 0DVDU\N

Each of the contributions was peer-reviewed by two anonymous reviewers in addition to the main reviewer prior to the publication of this volume.

FIRST EDITION

, 6 % 1
(online: iPDF; available at https://anglistika.upol.cz/olinco2021proceedings/)

DOI: 10.5507/ff.22.24461489

Table of Contents

Alphabetical List of Authors

Gábor Alberti
University of Pécs
3pFV +XQJDU\

Svitlana Antonyuk
University of Graz
Graz, Austria

Benedetta Baldi
University of Florence
Florence, Italy

0LFKDHOD ýDNiQ\RYi
Palacký University
Olomouc, Czech Republic

Luca Cilibrasi
Charles University in Prague
Prague, Czech Republic

Tamás Csontos
Károli Gáspár University
RI WKH 5HIRUPHG &KXUFK
%XGDSHVW +XQJDU\

Judit Farkas
University of Pécs
3pFV +XQJDU\

Volker Gast
Friedrich Schiller University
Jena, Germany

Pamela Goryczka
University of Vienna
Vienna, Austria

Dalina Kallulli
University of Vienna
Vienna, Austria

3UHGUDJ .RYDþHYLü
University of Novi Sad
Novi Sad, Serbia

Sabine Laszakovits
Austrian Academy of Sciences
Vienna, Austria;
University of Connecticut
Storrs, CT, USA

Chang Liu
Université Paris 8/UMR 7023 CNRS,
Paris, France;
Université de Picardie Jules Vernes,
Amiens, France

Markéta Malá
Charles University in Prague
Prague, Czech Republic

1LNROD 0DOHþNRYi
Charles University in Prague
Prague, Czech Republic

Tilda Neuberger
+XQJDULDQ 5HVHDUFK &HQWUH
%XGDSHVW +XQJDU\

Mark Newson
Eötvös Loránd University
%XGDSHVW +XQJDU\

Elisa Piccoli
& D ¶) R V F D U L 8 Q L Y H U V L W \ R I
Venice, Italy

Seid Tvica
Georg-August-University
Göttingen, Germany

Václav Jonáš Podlipský
Palacký University
Olomouc, Czech Republic

Lucia Vlášková
Masaryk University
% U Q R &] H F K 5 H S X E O L F

Natascha Pomino
8 Q L Y H U V L W \ R I : X S S H U W D O * H U P D Q \

Francesca Volpato
&D ¶) R V F D U L 8 Q L Y H U V L W \ R I
Venice, Italy

Zsolt Prohászka
University of Pécs
3 p F V + X Q J D U \

Eva-Maria Remberger
University of Vienna
Vienna, Austria

Leonardo M. Savoia
University of Florence
Florence, Italy

â i U N D â L P i þ N R Y i
Palacký University
Olomouc, Czech Republic

+ D Q D 6 W U D F K R R Y i
Masaryk University
% U Q R &] H F K 5 H S X E O L F

Krisztina Szécsényi
Eötvös Loránd University
% X G D S H V W + X Q J D U \

Anna Szeteli
University of Pécs
3 p F V + X Q J D U \

Acknowledgements

The editors are grateful to all those who have helped make the fifth volume of the Proceedings of the Olomouc Linguistics Colloquium (and the ninth volume in the whole series) a reality. Above all, as always, we would like to thank all the authors for both their enthusiastic participation in the conference and their cooperation in the time-consuming HGLWRULDO SURFHVV :H ZRXOG DOVR OLNH WR H[SUHV
who devotedly participated in the process of accepting and reviewing the papers for the conference and later another round of the peer-reviewing process for the proceedings.

)LQDOO\ ZH ZRXOG OLNH WR H[SUHVV JUDWLWXGH
Faculty of Arts of Palacký University, Olomouc, for their efforts related to the organization of the Olomouc Linguistics Colloquium (OLINCO), which took place online in -XQH :H JUHDWO\ DSSUHFLDWH WKH DVVL-VWDQFH
WLQNRYi 9iFODY -RQiã 3RGOLSVNê DQG âiUND âLPiþ
E\ 5RVDOLD &DOOH %RFDQHJUD 0LFKDHOD ýDNiQ\RYi
âiUND 'YR iNRYi /XFLH 3 LE\ORYi DQG +DOLQD =DZL
the conference during the global pandemic would have been an almost impossible task.

Markéta Janebová, Joseph Emonds, and Ludmila Veselovská

Introduction

The articles in this volume are based on papers and posters presented at the Olomouc Linguistics Conference (OLINCO) at Palacký University in the Czech Republic in June 2021. At this conference, papers combine analyses of language structure with generalizations about language use. The essays included here can be seen, we think, as a representative sample of the conference contributions; several of the papers were SUHVHQWHG DW WKH WKHPDWLF VHVVLRQV 7KH ¿UVW RUJDQL]HG E\ -RQDWKDQ %REDOMLN DQG 3DYHO &DK session can be found in the Part I of the volume (authored by Pamela Goryczka, 1DWDVFKD 3RPLQR DQG (YD 0DULD 5HPEHUJHU DQG %DOGL 7KH FRQYHQRUV RI WKH RWKHU WKHPDWLF V a Psycholinguistic and Cognitive Perspective, were Jan Chromý and Norbert Vanek. Three papers from this session can be found in Part III of this volume (Luca Cilibrasi, .DWH LQD +DVDORYi DQG $OåE WD %UDEFRYi (OLVD âLPiþNRYi DQG 9iFODY -RQiã 3RGOLSVNê

Part I. Explorations in Morphology

In the opening paper, Svitlana Antonyuk discusses the status of East Slavic languages ZLWK UHVSHFW WR WKH 5LFK $JUHHPHQW +\SRWKHVLV East Slavic, along with the rest of Slavic languages, is well within the scope of what is SUHGLFWHG E\ WKH ZHDN 5$+

In her contribution, 0LFKDHOD ýDNiQ\RYi considers two types of de nominals in Czech, analyzing one type as derived from the verbal stem and the other originating directly from the verbal root. These origins result in distinct properties and can account for the more verbal nature of the stem derived nominals and the more nominal nature of the root derived nominals.

The objectives of Pamela Goryczka ¶V SDSHU DUH WZRIROG IRU empirical support from Italian for span-conditioned theme (vowel) allomorphy; one the other hand, it aims to add a valuable contribution to the theoretical advancement of the framework of Distributed Morphology.

Dalina Kallulli and Sabine Laszakovits investigate the feature composition of VR FDOOHG PRUSKRORJLFDOO\ ³VLPSOH[´ ZK HOHPH RI GRXEO\ ¿OOHG FRPSOHPHQWL]HU HuHFWV REVHUY FODXVHV LQ %DYDULDQ DQG RWKHU 6RXWKHUQ *HUP

In their contribution, Natascha Pomino and Eva-Maria Remberger argue that not only Spanish, Italian etc., but also French regular verbs possess Theme Vowels and WKDW VRPH LUUHJXODULWLHV RI YHUEDO LQÀHFWLR Using a Spanning approach within the framework of Distributed Morphology, they show the accuracy of the Suppletion Generalization.

,Q WKLV VHFWLRQ Leonardo M. Savoia and Benedetta Baldi discuss WKH UHDOL]DWLRQ RI UG SHUVRQ REMHFW FOLWLFV GLDOHFWV ZKHUH WKH DX[LOLDU\ Dyalis involved. The authors question the authors investigate is the theoretical status of morphology: the idea is that the same FRPSXWDWLRQDO UXOHV RI V\QWD[DQG SDLU PHUJ EDVHG RQ DJUHHPHQW LQ 3 IHDWXUHV

Part II. Explorations in Syntax

In his contribution, Tamás Csontos focuses on multiple complement verbs in passive FRQVWUXFWLRQV LQ (QJOLVK +H SURYLGHV DQ H[SO dative construction and themes in the double object constructions are not passivized.

3 U H G U D J . R Y D Ć Ć H Ć Ć shows that Serbian psych verb SE anticausatives behave GLuHUHQWO\ IURP W\SLFDO DQWLDFXDVDWLYHV LQ FKDUDFWHULVWLF RI VWUXFWXUHV LQYROYLQJ DQ H[V

Chang Liu ¶V SDSHU H[DPLQHV WKH V\QWD[RI WKH YH formation of the Locative Inversion with an obligatory localiser phrase in Mandarin Chinese.

The joint paper by Mark Newson and Krisztina Szécsényi seeks to account for why some movements allow case change while others do not. The authors have argued that whether the case of a moved DP changes is dependent on both the case involved and the domain which contains it.

Zsolt Prohászka, Gábor Alberti, Anna Szeteli, and Judit Farkas have contributed a paper devoted to a comprehensive (i.e. syntactic, semantic, statistical) GHVFULSWLRQ RI WKH +XQJDULDQ (P Qóri, Gábor Alberti, Anna Szeteli, and Judit Farkas have also positioned the pronominal component of this family of constructions, which is XOWLPDWHO\ DQ LQGH¿QLWH SURQRXQ IRUPDOO\ LG +DVSHOPDWK¶V VHPDQWLF PDS RI LQGH¿QLWHV

In Seid Tvica ¶V SDSHU LW LV DUJXHG WKDW D XQLIRU proposed by Clemens and Coon (2018) does not straightforwardly account for verb-initial orders in Kaqchikel, as the VSO-VOS alternation appears to be syntactic, rather than post-syntactic. Two potential ways of deriving the VSO-VOS alternation are discussed, involving antisymmetric (with uniform leftward movement) and symmetric analyses.

Part III. Explorations in Language Use across Modalities

/XFD &LOLEUDVL .DWH LQD +DVDORYi DQG $OåE WD repetition in Czech-English bilinguals and show that the patterns observed are comparable to those found in monolinguals with a language impairment. This partly contradicts previous work that claimed that nonwords may be used to disentangle GLvFXOWLHV UHODWHG WR ELOLQJXDOLVP IURP GLvF

In his study, Volker Gast SURYLGHV DQ H[SORUDWRU\ DQDO\
a corpus of TV interviews. The hypothesis that eyebrow raises metaphorically signal
openness receives support from the data, especially with respect to their occurrence in
WKH FRQWH[W RI HSLVWHPLF PRGDO H[SUHVVLRQV D
1LNROD 0DOHþNRYi DQG 0DUNpWD 0DOi H[SORUH
adverbsabsolutely and totally LQ SUHVHQW GD\ LQIRUPDO VSRNH
paper shows that they both appear to be following the same trajectory of change from
DQ LQWHQVL¿HU YLD D VWDQFH DGYHUELDO WR D GL
FXUUHQWO\ DW GLuHUHQW VWDJHV RI WKH SURFHVV
Tilda Neuberger ¶V VWXG\ DLPV WR LQYHVWLJDWH WKH
WKH +XQJDULDQ VLQJOHWRQ DQG JHPLQDWH YRLFHO
relationship between the acoustic and perceptual domains and shed light on the primary/
VHFRQGDU\ DFRXVWLF IHDWXUHV RI FRQVRQDQW OHQ
Elisa Piccoli and Francesca Volpato investigate oblique relative clauses,
structures typical of formal registers, in a group of Italian-speaking adolescents with
GHYHORSPHQWDO G\VOH[LD FRPSDUHG WR D JURXS
peers.
âiUND âLPiþNRYi DQG 9iFODY -RQiã 3RGOLSVNê ¶
H[SORUHV SKRQHWLF HuHFWV RI VZLWFKLQJ EHWZHH
ELOLQJXDOV KLJKO\ SUR¿FLHQW LQ D IRUHLJQ ODQJ
induced via elicited code switching and language switching in a picture naming task.
In the concluding paper, based on data from Czech Sign Language,
6WUDFKR RYi DQG /XFLH 9RMWČKRYi SKRQRORJLFDO +
(Sandler 2006) by proposing solutions to a number of theoretical problems. Moreover,
WKH DXWKRUV SUHVHQW D OH[LFRJUDSKLF DSSOLFDW
&]HFK 6LJQ /DQJXDJH OH[HPHV LQWR YDULDQWV DQG

:H KRSH WKDW UHDGHUV ZLOO ILQG WKH SDSHUV LQFO
and their fellow researchers. It was both challenging and gratifying to organize and par-
ticipate in the conference, in 2021 more than ever because of the globac pandemic, and
QRZ ZH ZDQW WR H[WHQG WKH FKDOOHQJHV DQG WKH
audience of those who can participate via the written word, which was, as we tradition-
ally like to say at this place by way of conclusion, invented by our species so that the
SOHDVXUHV DQG EHQHILWV RI ODQJXDJH XVH FRXOG E

Markéta Janebová and Joseph Emonds

Part I. Explorations in Morphology

Object-Shifting and Head-Raising One's Way to Discourse Configurationality

Svitlana Antonyuk

University of Graz, Graz, Austria

svitlana.antonyuk@uni-graz.at

Abstract: I discuss the question of the status of East Slavic languages with respect to the Rich Agreement Hypothesis (henceforth the RAH), arguing that despite appearances, East Slavic, along with the rest of Slavic languages, is well within the scope of what is predicted by the weak RAH. This conclusion relies on a particular conception of head raising, namely that developed in Roberts (2010), i.e., syntactic raising as Defective Goal incorporation. To the extent the analysis proposed here is successful, it provides further support for Roberts (2010) while removing Slavic languages from the list of challenges for the RAH. Overall, two types of head raising are posited in East Slavic Russian and Ukrainian: obligatory V-to-Asp and optional raising above Asp/into the T domain, and both are treated as syntactic in nature.

Keywords: head movement; East Slavic; Rich Agreement Hypothesis; optionality; discourse configurational languages

1. Introduction

There is an important well-known observation about the existence of a correlation between the complexity of verbal (inflectional) morphology and word order in a language according to which the presence of rich agreement morphology positively correlates with the raised/vP-external position of the verb and specifically with the verb being located in T (Rohrbacher 1999; Vikner 1997; Bobaljik 2002 i.a.):

(1) (a) þú veist [að ég skil alls ekki japönsku].
you know that I understand at.all not Japanese
"You know that I don't understand Japanese at all." Icelandic

(b) Du vet [att jag inte alls förstår japanska].
you know that I not at.all understand Japanese
"You know that I don't understand Japanese at all." Swedish

In the above examples, due to Holmberg and Roberts (2013), a morphologically rich language, Icelandic (1a), has the lexical verb in the obligatorily raised position (as indicated by its placement to the left of the adverb and negation) whereas in Swedish (1b), the verb has not raised, occurring to the right of these vP-edge-marking elements and is thus widely regarded to be inside the vP in such cases. The same well-known pattern is found in French (2a), where a temporal adverb marks the verb's obligatorily raised position (Pollock 1989), compared to (Modern) English (2b), in which the verb must follow the adverb and is believed to remain vP-internal.

(2) (a) Jean embrasse souvent Marie. French
John kisses often Mary
"Jean often kisses Marie"

(b) John often kisses Mary English

Again, there is a correlation in these and similar cases between the richness of verbal agreement morphology and verb raising: while French exhibits both, present day English lacks both (see Roberts 1993; Holmberg and Roberts 2013). There is a debate in the literature on whether a strong (bi-conditional) or a weak (unidirectional) version of the Rich Agreement Hypothesis (3) is correct, the exact nature of morphology that is relevant to the RAH as well as the directionality of the correlation, i.e., whether it is richness of morphology that is responsible for rich functional structure or whether it is rich(er) functional structure that drives movement.[1]

(3) Move V to T iff T has rich inflection (RAH, the strong version)

Whichever view of the RAH and the related issues one takes, however, one thing is immediately apparent: Slavic languages, known for their morphological richness, especially

1 See Bobaljik and Thrainsson (1998); Bobaljik (2002) i.a. on the untenability of the strong RAH and the directionality going from rich functional structure to rich morphology to V-to-T.

richness of verbal agreement morphology, are predicted by the RAH to exhibit obligatory verb raising to Tense of the kind found in Icelandic and French. As is well-known, however, Slavic languages such as Russian, Ukrainian, B/C/S or Polish do not in fact exhibit verb raising, thus posing a significant challenge even for the weak version of the RAH (see Koeneman and Zeijlstra 2014, henceforth K&Z). In section 2 of the paper I review the data from Russian as presented in the literature and argue (contra K&Z) that the actual empirical picture is actually worse than the one they argue against. Nevertheless, in section 3 I argue that one view of verb raising, namely that in Roberts (2010), coupled with certain assumptions about functional structure and featural content of (East) Slavic (henceforth ES) presents a straightforward solution to this problem, one that places Slavic languages firmly within the purview of the weak RAH. Furthermore, I argue that the account captures the peculiar distribution of head movement in (E)S and briefly discuss some of the predictions. Section 4 presents my conclusions.

2. Head Movement in East Slavic: The Empirical Picture

As far as the status of verb raising in Russian is concerned, the existing syntactic literature appears to be in almost complete agreement that the verb in Russian does not undergo V-to-T (Bailyn 1995; 2012; cf. King 1993; see Dyakonova 2009 and Gribanova 2017 on verb raising into AspP[2]). Thus, sentences such as (4a–b) from Bailyn (2005) show the canonical placement of the verb, its position to the right of low manner or frequency adverbs generally taken to indicate it remains vP-internal.

(4) (a) My vnimatel’no S U R þ L W D O L pravila. Russian
we carefully ‰ ~ ‹ • read rules
“We have carefully read the rules.”

(b) My þ D V W R þ L W D O L pravila.
we often , † ‰ read rules
“We read the rules often.”

2.1 Koeneman and Zeijlstra (2014)

Such facts are of course highly problematic even for the weak RAH, presenting direct counterevidence to the prediction that a language with rich verbal agreement morphology

2 Bailyn (2004) argues that the only context in which the Russian verb undergoes V-to-T of the French kind is OVS sentences. This conclusion has been heavily criticized (see Slioussar 2007; 2011 i.a.) and largely abandoned, though Antonyuk (2021) argues that the verb in OVS can indeed appear in a raised position; crucially though, the raised position of the verb in such cases is argued to correlate with the verb’s status as given/D-linked material. Some of this evidence is reviewed in section 2.3.

will exhibit mandatory V-to-T, suggesting that the RAH might be nothing more than a correlation. Recognizing the problem, K&Z set out to reanalyze the above Russian data and argue that (4) is in fact fully compatible with the predictions of the weak RAH. Specifically, they argue that Russian is a language with obligatory high attachment of manner/frequency adverbs, thus the typically assumed adjunction option in (5a) is argued to be ungrammatical, with the sentences in (4) having structure roughly as in (5b), with a manner adverb *vnimatel'no/attentively* attached to TP.

(5) (a) [$_{IP}$ my [$_{vP}$ vnimatel'no S U R þ L W Dv D O D]]

(b) [$_{FP}$ my [$_{IP}$ vnimatel'no [$_{IP}$ S U R þ L W [$_{vP}$ t$_{V}$ [$_{VfIn}$ þ D O D]]]]]

The evidence for the above claim comes from sentences involving sentential negation, which is proclitic to the verb in Russian, with NEG moving to C together with the verb in imperatives (6a) and interrogatives (6b), schematized in (7):

(6) (a) Ne pey vodku þ D V W R
‡ ~ € drink vodka often
³ ' R Q R W G U L Q N Y R G N D R I W H Q ´

(b) Ne p'eš li ty vodku þ D V W R "
‡ ~ € drink Š you vodka often
³ ' R Q ¶ W \ R X R I W H Q G U L Q N Y R G N D " ´

(7) [CP [ne-V$_i$]$_j$ [NegP t$_j$ [vP t$_i$]]]

As argued by K&Z, the vP-attachment site view of manner/frequency adverbs predicts that sentences with sentential negation such as (8) will be grammatical, which they claim is contrary to fact, whereas high-attachment view correctly entails the ungrammaticality of (8) and grammaticality of (9) and (10), the latter case involving adverb TP-attached on the right.[3]

(8) (a) " " 7 \ ne S U R þ L Dv D O D pravila.
you ‡ ~ € ‰ ~ read carefully rules
"You haven't read the rules carefully."

3 The data in (8)–(10) are given here with K&Z's grammaticality judgments.

(b) " " 7 \ ne þ L W D O þ D V W R pravila.
you ‡ ~ € , † ‰ read often rules
“You haven’t read the rules often.”

(9) (a) Vnimatel’no ty ne S U R þ L W D O pravila.
carefully you ‡ ~ € ‰ ~ read rules
“You haven’t carefully read the rules.”

(b) ý D V W R ty ne S U R þ L W D O pravila.
often you ‡ ~ € , † ‰ read rules
“You haven’t often read the rules.”

(10) (a) Ty ne S U R þ L W D O D vnimatel’no.
you ‡ ~ € ‰ ~ read rules carefully
“You haven’t carefully read the rules.”

(b) Ty ne þ L W D O pravila þ D V W R
you ‡ ~ € , † ‰ read rules often
“You haven’t often read the rules.”

On the basis of these data K&Z conclude that adverbs in Russian must be attached higher than vP or involve right-adjunction, therefore low adverbs cannot diagnose verb movement and Russian thus fails to provide evidence against the weak RAH.

2.2 Another Look at East Slavic Head Raising Possibilities

It should be noted, however, that the empirical data in (8)–(10) that K&Z’s account of Russian verb raising is based on is far from straightforward. To my ear, the sentences in (8) are in fact the most neutral way of expressing sentential negation, which means that the vP-attachment of low adverbs is not only possible, but is in fact preferred. Furthermore, if low adverbs are vP-attached and negation is generated left of the adverbs (as is minimally suggested by the neutral word order in (8), but see also Gribanova 2017) the natural conclusion then is that in sentences with sentential negation the verb raises from vP into NEG head, with any further movement then taking place with negation procliticized to the verb, effectively as schematized in K&Z’s (7). Examples in (9), on the other hand, are highly marked, and are only acceptable in a particular non-neutral prosody (where the sentence-initial adverb and verb carry the strongest rising and falling pitch accents respectively), which suggests that this attachment option, if available to speakers, is marked compared to (8). Finally, the examples in (10) have a more likely derivation: they can be straightforwardly derived by object raising to the left of the adverb, thus again implicating the verb’s vP-external (raised into NEG) position in such cases.

In the closely related Ukrainian, where the verb raising options appear to be the same as in Russian, Object Shift (Middle Object Scrambling in Mykhaylyk 2011) targets the position to the left of the adverb, which obligatorily results in semantic effects for the shifted XP (specificity/partitivity interpretation, see also Antonyuk and Mykhaylyk 2022). The default assumption of parallelism between the two languages as far as object raising possibilities are concerned thus suggests a way to test the hypothesis about the derivation of (10): if the order in (10) is derived via verb and object vacating the vP (rather than via adverb TP-right adjunction), the object in (10) will have to be obligatorily interpreted as specific/carrying an existence presupposition. The prediction is, indeed, correct.

Two additional observations are relevant here. As argued in Antonyuk (2021), on the analysis of quantifier float (Q-float) as left-behind of A-movement, with a number of landing sites available to it, both vP-internally and vP-externally. The data is thus simply incompatible with an obligatory TP-adjunction of low adverbs. To see why, consider (11), which indicates the canonical vP attachment of low adverbs:

(11) Maks (vse) korobki s domašnej utvar'ju Russian
Max (all) boxes with home furnishings
(vse) medlenno (vse) > V O R å (vse) v mašinu (vse)].
(all) slowly (all) put (all) in car (all)
"Max put in the car all the boxes with home furnishings"

First, in a sentence with a heavy object as in (11), a sentence-final adverb placement would be marginally possible, but a placement immediately following the subject is not merely dispreferred, it is highly unnatural. Furthermore, attempting to nevertheless analyze the order in (11) as reflecting TP-attachment of the adverb and the verb raised into T would suggest that the subject and the object are higher than TP, which implicates A-bar positions. However, there is plenty of evidence to suggest that the subject in both SVO and SOV sentences is in its canonical A-position, Spec,TP, in ES (see Bailyn 2012 on Russian; Antonyuk and Mykhaylyk 2022 on Ukrainian), hence the order in (11) is incompatible with obligatory TP-attachment of low adverbs. I conclude that the position of the verb in ES is only compatible with it being vP-internal when preceded by a adverb and vP-external when preceding the adverb.

2.3 Clarifying the Empirical Domain

Nevertheless, to draw the conclusion that the verb in ES never raises outside the vP would be an oversimplification. We have seen that the verb does raise across an adverb in sentences with sentential negation such as (8) and (10). Furthermore, there is evidence that the verb in ES can in fact occur vP-externally even when sentential negation is not involved, and, presumably, these cases involve verb raising into the Tense domain.

Consider the following data from Ukrainian (the hashtag symbol indicates infelicity of the examples, rather than ungrammaticality[4]):

(12) (a) â þ R Marijka robyt’] U D Q N X " Ukrainian
What Mary does in.morning
“ : K D W G R H V 0 D U \ G R’ L Q W K H P R U Q L Q J "

(b) Marijka švydko gotuje #(švydko) jaješnju
Mary.NOM quickly cooks (quickly) scrambled eggs.ACC
i E L å \ W ¶ na robotu.
and runs on work
“Mary quickly cooks scrambled eggs and hurries off to work”

(13) (a) â þ R P R å H Marijka P R å H prygotuvaty ã Y \ G N R "
What can Mary can cook quickly
“ : K D W F D Q 0 D U \ F R R N T X L F N O \ "

(b) Marijka ##(švydko) gotuje švydko lyše jaješnju.
Mary. ‡ ˆ † (quickly) cooks quickly only scram. eggs. z | |
“Mary cooks only scrambled eggs quickly”

(14) (a) â þ R vidomo pro tsju N Y L W N X "
What known about this flower
“ : K D W L V N Q R Z Q D E R X W W K L V I O R Z H U " ´

(b) Tsju kvitku dobre znaly ##(dobre) drevni greky.
This flower well knew (well) ancient Greeks ‡ ˆ †
“This flower was well known to Ancient Greeks”

(c) ' U H Y Q L Greky dobre znaly tsju kvitku.
Ancient Greeks well knew this flower. z | |
“Mary cooks only scrambled eggs quickly”

(15) (a) Xto znaje ã þ R V ¶ pro tsju N Y L W N X "
Who knows something about this flower. z | |
“ : K R N Q R Z V D Q \ W K L Q J D E R X W W K L V I O R Z H U " ´

4 The Russian counterpart of (14b) is due to Natalia Slioussar (p.c.)

(b) Tsju kvitku ##(dobre) znaly dobre lyše drevni greky.
This flower (well) knew well only ancient Greeks
"Ancient Greeks were the only ones to know this flower well"

The data above demonstrate the following descriptive generalization, which holds true of many languages cross-linguistically as well as of all Slavic languages: the word order in Slavic is (re)arranged in such a way as to place given/D-linked material before new elements (see esp. Titov 2007, 2011, 2012). Thus, (13b), for instance, shows that the verb can raise to the left of low adverbs; however, this order is fully determined by the verb's Information Structural (IS) status and is not as acceptable in (12b), where the verb represents new information. Another observation is that the lexical verb in ES tends to raise outside the vP mostly in OVS clauses, that is, there appears to be a positive correlation between the givenness/vP-external position of the object on the one hand and the D-linking/vP-external position of the verb on the other. A final point to note is that most of the above orders that are non-neutral are nevertheless not ungrammatical, merely dispreferred (marked with hashtag signs); and even those cases that are very strongly degraded (on neutral intonation) can be made fully acceptable via prosodic recontouring, that is, via a change in the prosodic realization of the sentence which acts as an alternative means of signaling the intended semantics (see Antonyuk and Mykhaylyk 2013 for an experimental investigation). In the context of our discussion, coupled with our earlier conclusions, this yields the following broad empirical generalization to account for:

(16) the verb in ES does not undergo V-to-T but is able to raise into the Tense domain for IS reasons. This raising is optional, with prosodic recontouring providing an alternative means of encoding IS-relevant features (givenness/D-linking).

While this generalization is not entirely new, it is almost never acknowledged in studies focused on the status of Slavic languages with respect to the RAH, portraying Russian, for instance, as either patterning with English in not allowing V-to-T (Bailyn 1995) or with French in requiring it (Koeneman and Zeijlstra 2014). Furthermore, generative literature on Russian mostly treats this empirical situation as clear evidence of the phonological nature of head movement in Russian, which reflects mapping to IS (Bailyn 1995; Kallestinova and Slabakova 2008; Slioussar 2007 i.a.). In this paper and in Antonyuk (in preparation) I argue that this empirical picture, which extends to all of Slavic, shows that instead of a two-way typological split, i.e., (Modern English, Swedish, Russian) vs (Early Modern English, Icelandic, French), there is a three-way split, with most Slavic languages representing the third group, where verb raising into the Tense domain is possible, but never obligatory. The challenge for the RAH, of course, at this point appears to be all but insurmountable, since such an empirical profile of ES

effectively reduces the otherwise fairly robust typological generalization to a correlation that holds robustly in select languages, at best.

In the next section I will argue that one conception of head raising, that of Roberts (2010), allows for a natural inclusion of Slavic languages within the purview of the RAH. Specifically, I will argue that the verb in ES undergoes head raising to Asp and no further for reasons that are fully compatible with the RAH, tracing the differences between ES on the one hand and English, Icelandic, French, etc., on the other to differences in the richness of functional projections (Bobaljik and Thrainsson 1998; Bobaljiik 2002 i.a.) and to Defective Tense in ES which precludes V-to-T, allowing only V-to-Asp. Finally, I will elucidate briefly how this account deals with the optional head movement beyond AspP observed in the data.

3. Head Raising in East Slavic: The Account

3.1 Roberts (2010) on V-to-T as Incorporation

Roberts (2010) proposes a reanalysis of head raising as Goal incorporation, where features of a Defective Goal, defined as in (17), copied onto the Probe upon the AGREE relation are pronounced on the Probe, which for all intents and purposes is indistinguishable from movement.

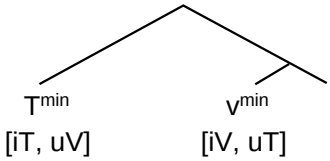
(17) Defective Goal: a Goal G is defective iff G's formal features are a proper subset of those of G's Probe P. (Roberts 2010, 62)

Consider the ingredients implicated in the French V-to-T, for example. According to Roberts, T in French has an interpretable T feature (iT) and an uninterpretable V feature (uV), while V has the opposite: an interpretable V feature (iV) and an uninterpretable T feature (uT):

---

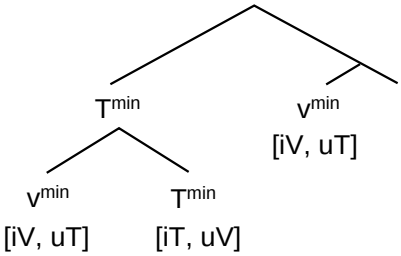
5 This conclusion crucially depends on the assumption that head raising in Slavic is fully comparable in this regard to head movement in other languages, both those with and without 9 WR 7 6SHFL¿FDOO\ LI KHDG UDLVLQJ LQ (6 LV WUHDW apparently optional character, then the obligatoriness of head raising in French would provide a reason to treat it as syntactic movement. However, such an approach is not sustainable; it also leaves no hope of providing a uniform account for head raising from a crosslinguistic, typological perspective. The assumptions that (i) head raising is syntactic and (ii) languages with DQG ZLWKRXW 9 WR 7 GR QRW IXQGDPHQWDOO\ GLuHU LQ assumptions adopted in this paper. See Antonyuk and Mykhaylyk (2022); Matushansky (2006); 5REHUWV L D IRU YDULRXV W\SHV RI HYLGHQFH VXSS for a comprehensive overview of various approaches to head movement and their challenges; see GHQ 'LNNHQ DQG 'pNiQ\ RQ WKH DSSOLFDWLRQ RI 5RE to the typology of clitics and noun incorporation.

(18) AGREE configuration L J Q R U L Q J W K H 3 I H D W X U H V L Q 7



Upon AGREE, V's features get copied into T and are thus now present in two places:

(19) French and other V-to-T languages upon AGREE:



Since T c-commands V, the V feature in T and the V feature in V now form a chain; upon linearization, the head of the chain (i.e., V features in T) are pronounced, with the tail of the chain silenced, thus giving the impression of verb raising into T. On this conception of head raising, the difference between French and Modern English comes down to the fact that T in English does not have a V feature, therefore the output of AGREE cannot be head incorporation, resulting in lack of V-to-T.

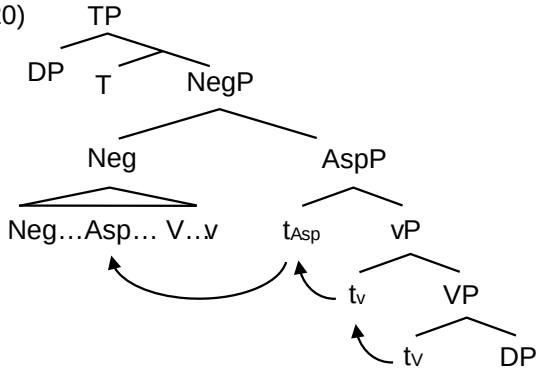
3.2 Roberts (2010) Applied to East Slavic

To derive the ES facts, we need to adopt the above mechanism from Roberts (2010) coupled with several independent assumptions about Slavic. Let us first consider the logic of this account and what one would need to derive the absence of V-to-T in ES. A straightforward view of the differences between English, French and ES in the context of Roberts (2010) is that the ES verb must have a feature that T lacks, to ensure that the features in V are not in a proper subset relation to the features in T. The crucial ques W L R Q L V Z K D W F R X O G W K L V I H D W X U H E H " , E H O L H Y H the most prominent morphological property of the Slavic verb – its aspectual system,

the relevant feature thus having to do with Aspect, and the presence of AspP would ensure that v/V cannot agree with T directly, having to agree with Asp[6] first.

I will assume, following Dyakonova (2007) and Gribanova (2017) on Russian that ES languages have an intermediate functional projection between vP and TP, namely AspP, yielding the structure in (20), as defended in Gribanova (2017). (20) shows the lexical verb raising into Asp, followed by raising into NEG, schematically representing the examples such as (8)/(10).[7] The claim is thus that the verb undergoes head raising/ incorporation into Asp, followed by raising into Neg (in cases of sentential negation), but not higher, thus differing from French in this regard.[8]

(20)



It is crucial for this account to not only assume that ES languages have an intermediate projection AspP as schematized above, but also that French, for instance, does not. As it happens, there is independent evidence to suggest that this is indeed so. It is well known

---

7 KLV LV VWULFWO\ VSHDNLQJ QRW VR $V SRLQWHG RX qua incorporation mechanism allows for a situation where the Probe and the Goal of AGREE are not in a local relation, thus the mere presence of AspP does not by itself preclude AGREE between T and V from taking place. However, I will argue further that V in Slavic must contain DQ[X]QLQWHUSUHWDEOH $VS IHDWXUH ZKLFK PHDQV WKDW L before it can agree with T, thus deriving the correct result.

7 I further assume, following Dyakonova (2009), that the manner/frequency adverbs attach to AspP rather than vP. See Dyakonova (2009) for arguments in support for this view.

8 The optional movement into T discussed earlier is thus treated here not as optional V-to-T but as syntactic movement with IS-related consequences that feeds on the output of head raising qua incorporation, its optional character following naturally from the fact that the same semantics can be encoded prosodically. I thus assume with Roberts (2010) that in addition to incorporation, ZKLFK LV PRVW\ GHWDLOHG LQ 6KDHQRPHQRQ WKHUH FDQ EH R PRYHPHQW 6HH DOVR WKH GLVFXVVLRQ LQ 'pNiQ\ RQ

that Slavic aspectual perfective morphology encodes certain semantic properties: Russian bare plurals and bare mass nouns are interpreted as definite or specific in the context of a perfective verb, whereas they can be interpreted as definite or indefinite in the context of an imperfective verb (see de Swart 2012 for an overview):

(11) (a) Petja þ L W D O stat'i/literaturu Russian
Peter read. , † ‰ ‰ z Œ • Œ articles/the literaturez | |
"Peter was reading articles/the articles/literature/the literature"

(b) Petja S U R þ L W D O stat'i/literaturu
Peter ‰ ~ ‹ •-read- ‰ z Œ • Œ articles/the literaturez | |
"Peter read the articles" (Borik 2002)

Whether Slavic prefixes should be treated as perfectivity markers (Smith 1991/1997, Borik 2002 for Russian) or whether they are telicity inducing, as is the case in (21b) D E R Y H V H H) L O L S I R U &] H F K $ U V H Q L M H Y L ü I I Swart (2012), grammatical aspect and aspectual class are clearly intertwined in Russian but (crucially for our purposes), separate from Tense. The idea that Aspect may be what is responsible for the additional features in V in Russian and other Slavic languages appears to be further supported by the fact that aspect in Romance not only does not encode the same definiteness/specificity distinctions as do Slavic languages, but French and Romance languages more generally exhibit fusing of past tense and perfective/ L P S H U I H F W L Y H D V S H F W X D O P R U S K R O R J \ & R Q V L G H U W Swart (2012), which demonstrate the impossibility of mapping (22a) to a compositional structure in which the past tense and the perfective aspectual operators are separate. Such differences between Slavic and Romance languages are thus crucial for my account, implicating a separation of Tense and Aspect into independent functional projections in East Slavic but their fusion into a single functional projection in French.

(22) (a) Il écrivit sa tèse en 2009. French
He wrote ‰ his thesis in 2009
"He wrote his thesis in 2009."

(b) [past [perf [he write his thesis]]]

While the above discussion about Slavic Aspect encoding semantic features such as definiteness and specificity makes it tempting to propose that the feature that Asp and V heads share to the exclusion of T has to do with definiteness/specificity semantics, I believe this step would be an overcomplication and that it is sufficient for us to assume that V hosts an uninterpretable Asp feature (uAsp), while an interpretable Asp feature

(iAsp) is hosted in the Asp head, thereby modeling this proposal exactly on the reasoning in Roberts (2010). Upon establishing AGREE between Asp and V, with respect to which V acts as a Deficient Goal, the V feature in Asp will be pronounced as the head of the chain, whereas the V feature in V will be silenced as the tail of the chain.[9] Crucially though, T must not have the Asp feature, thus precluding incorporation of Asp into T due to Asp not being a Deficient Goal with respect to T.

3.3 A Simpler Alternative (No-Tense-Projection) Account?

At this point we should address typological research that posits the absence of the Tense SURMHFWLRQ LQ WKRVH 6ODYLF ODQJXDJHV WKDW O 0LJGDOVNL -XQJ DQG 0LJGDOVNL 7 appears to suggest a much simpler take on the lack of V-to-T in EST.[10] The absence of the Tense projection in (E)S would derive the lack of V-to-T quite trivially: there can be no head raising into T if there is no Tense projection, hence no T to probe V. In order to preserve the account developed here, one might thus be tempted to disregard the above research and assume the universality of TenseP. I believe taking this position would not be beneficial as it would lead to overlooking a wealth of relevant typological generalizations and insights. In fact, I argue that assuming the above accounts must be fundamentally correct can help us strengthen the present account and extend its coverage to all of Slavic.

Interestingly, the no-Tense-projection accounts do not suggest that the functional Tense layer is absent in all of Slavic: in fact, Bulgarian and Macedonian are argued to be different from the rest of Slavic in exhibiting robust Tense morphology and straightforward Tense-Aspect separation and are therefore argued to project Tense, in contrast to languages with poor/residual tense morphology such as Russian/Ukrainian/BCS/ HWF VHH HVS 0LJGDOVNL.[11] Crucially, Bulgarian is also said to have V-to-T.[12] This, of course, appears to falsify the present account and provide strong support for the no-Tense-projection accounts in terms of the reason for the lack RI 9 WR 7 , EHOLHYH WKHUH LV DQRWKHU ZD\ WR LQV that "the absence of morphological realization is the reflex of structural deficiency in

---

9 This discussion assumes that the same AGREE-based incorporation has already taken place between v and V, thus the features of V are already in v at the time Asp and v undergo AGREE. I continue to refer to V rather than v for ease of exposition.

10 7KH UHOHYDQFH RI WKLV YLHZ RI (6 ZDV SRLQWHG RXW

11 Macedonian is reported to be less straightforward, having started the process of losing Tense-Aspect distinctions that Bulgarian still makes. In what follows I therefore refer to Bulgarian alone.

12 See Krapova (1999) on the Bulgarian auxiliary system and Harizanov (2019) for a recent overview of the literature as well as an argument, based on Bulgarian participle fronting, for the XQL¿FDWLRQ RI KHDG PRYHPHQW ZLWK ;3 PRYHPHQW

WHUPV RI WKH DEVHQFH RI D SDUWLFXODU SURMHF
I propose that the ES vs Bulgarian contrast, viewed in the context of Roberts' (2010) proposal, provides us with another perspective on what structural deficiency could mean. Specifically, Ipropose that structural deficiency of Tense in most of Slavic (as diagnosed LQ 7RGRUYLü EH XQGHUVWRRG LQ WHUPV RI IH

Consider the details of this proposal. The difference between Russian and Ukrainian on the one hand and Bulgarian on another cannot reside in the presence/absence of the AspP as I have argued to be the case with French, as the above research makes a strong case for well-developed Tense and Aspect systems in Bulgarian, yet Bulgarian has V-to-T and Russian and Ukrainian do not.[13] I propose the key difference is that the ES Tense itself is deficient in lacking the uAsp feature while the Bulgarian T head is not and does have the uAsp feature. Consider what this gives us: if Bulgarian T has a uAsp feature, then the Asp head will count as a deficient Goal wrt T (assuming, of course, with Roberts, that T always has other features) and Asp incorporation into T will take place in Bulgarian in the same way that V-to-T takes place in French. The result we have derived is that the absence of V-to-T in ES is not merely due to the presence of an intermediate projection between T and v and the presence of a corresponding feature in v/V. Crucially, for the Goal not to count as deficient with respect to the Tense Probe, Tense must itself be deficient in lacking the corresponding feature. The account of ES developed here also illustrates DQRWKHU SURSHUWY RI 5REHUWV¶ DFFRXQW SLQSRLQW of Goal deficiency is a relative rather than an inherent one: a Goal can be deficient with respect to one Probe (here, T) yet non-deficient with respect to another (here, Asp). The elaboration of the ES vs Bulgarian case further illustrates an important related point: a Goal is non-deficient when it has features that are not also present in the Probe, which entails that whenever a Goal is non-deficient, a Probe must be.[14,15] I conclude that as far as V-to-T phenomena are concerned, recasting the no-TenseP position in terms of featural deficiency of TenseP appears quite promising.[16]

---

13 Note that B/C/S, Slovenian, Polish and other Slavic languages pattern with ES in this regard and in terms of the lack of V-to-T. Thus, the proposal made here for Ukrainian and Russian extends naturally to these languages.

14 :KLOH WKHVH QRWLRQV DUH FOHDUO\ LQWHUUHODWHG GLuHUHQW HYHQ RI SDUWLFXODU LWHPV RQ RQH DQDO LW KDV PRUH IH ZKLOH %UR DE LV GH¿FLHQW ZKHQ LW VKDUHV D ZKROH IHDWXUH RI WKH 3UREH (DVW 6ODYLF

15 7HQVH GHH¿FLHQF\ FDQ EH PRUH RU OHVV GUDVWLF RQ WKH SDW LQ (DVW 6ODYLF HVS /DYLQH DQG)UHGLQ RQ 'HFHPEHU 7HQVH :KLOH FR DUH VWLOO FRQQHFWHG E\ WKH LGHD WKDW 7HQVH ODFNLQJ F

16 It remains to be seen to what extent the phenomena accounted for in the no-TenseP tradition VHH HVS %RãNRYLü FDQ EH FDSWXUHG LQ WKH 'H¿FL

3.4 On Optional IS-Related V Raising in (East) Slavic

Let us now return to the problematic aspect of East Slavic verb raising into the T domain, namely its apparently optional character. I believe the empirical facts of ES demonstrate an intuition pursued in a number of recent accounts on head movement, most prominently Matushansky (2006), Roberts (2010), Harizanov and Gribanova (2019), namely the idea that what is referred to as 'head movement' likely subsumes more than one operation. Thus, I propose that ES languages have V-to-Asp, which, on Roberts' (2010) account assumed here can only be understood as a syntactic operation (with AGREE at its core). What we observe in East Slavic OVS and other contexts where the verb appears to optionally raise into the Tense domain when it is given/D-linked, is, I argue, syntactic movement as well, though distinct from the syntactic operation implicated in V-to-Asp.

While numerous accounts have treated such movement as post-syntactic, not in the least because of its optional and (presumably) syntactically vacuous character (though see esp. Matushansky 2006 on this latter point), I believe there is important overlooked evidence which implicates the syntactic nature of optional IS-related verb movement in East Slavic, namely its similarity to Object Shift in Ukrainian. As shown in Antonyuk and Mykhaylyk (2013), Ukrainian Object Shift exhibits the same type of apparent optionality described for IS-related verb movement earlier in this paper. Specifically, Object Shift, which is obligatory in situations where contextual information implicates the specific/partitive semantics of the object is nevertheless optional in that the same semantic interpretation can be obtained without syntactic movement via prosodic recontouring. That is, encoding the relevant semantics is obligatory, what is optional is merely the means of encoding, i.e., via syntactic movement or via prosody. Now, Ukrainian Object Shift is, beyond doubt, a syntactic operation[17], the fact that the exact same pattern of movement/prosodic encoding alternation is also observed with verb raising then strongly suggests that the verb movement in question is also syntactic. Thus, while length considerations prevent me from elucidating many of the relevant details, if the account broadly sketched here is on the right track, it achieves several things: (i) removes Slavic languages from the list of challenges to the weak RAH; (ii) provides further crosslinguistic support for Roberts' (2010) treatment of head raising; (iii) relates typological differences wrt V-to-T to differences in the functional sequence (Bobaljik and Thrainsson 1998; Bobaljik 2002) and featural specification of both Probe and Goal of AGREE relation (iv) identifies two types of head raising in ES (V-to-Asp and IS-related movement) and (v) argues for the purely syntactic character of both types of movement.[18]

17 See esp. Antonyuk and Mykhaylyk (2013) on the interaction of Object Shift and Negation in Ukrainian.

18 See Antonyuk (in preparation) for a larger cross-Slavic proposal encompassing the issues discussed here.

3.5 The Predictions

The account proposed here makes a number of testable predictions that cannot be properly explored here for space reasons. As an illustration, however, consider the following. The idea that differences wrt V-to-T are to be traced to differences in the functional sequence and the featural specification of heads suggests that there will likely be other significant differences between the languages under investigation that are similarly traceable to the presence/absence of AspP unbundled from the TenseP. I suggest this is indeed so and the numerous observable differences manifest themselves in the expected direction. For instance, as is well-known, constructionist approaches to argument structure face challenges due to observed limitations on syntactic malleability of roots (see Ramchand 2008; 2013 i.a.). Thus, theories that attribute no idiosyncratic lexical meaning to verbs beyond the encyclopedic meaning of roots (e.g., Borer 2005) predict greater syntactic freedom on the part of roots than what is actually observed. For instance, (23) in English on such theories is predicted to be a grammatical sentence, contrary to fact:

(23) 0 D U \ slept the baby. English

Slavic aspectual morphology has been known to interact with argument structure in numerous non-trivial ways, e.g., by contributing an argument, adding a Result phrase or by creating whole argument structure alternations such as the Spray-Load alternation

$ Q W R Q \ X N D E $ U V H Q L M H Y L ü D- E 4 X D J
Q L X V 7 D W H Y R V R Y ä D X F H U L D $ V G L V F X V

for instance, the counterpart of (23) is perfectly grammatical in Ukrainian and Russian, with aspectual morphology providing the mechanism needed for coaxing the root ‘sleep’ into an (obligatorily) transitive frame taking an optional Inanimate Causer argument:

(24) Marijka pry-spa-l-a dytynu
Mary. ‡ ˆ † z Œ ‰-sleep- ‰ Œ • baby. z Œ € ‡
(svojim spivom /pisneju) Ukrainian
self. , ‡ Œ • < signing. , ‡ Œ • < song.i ‡ Œ • <
“Mary put the baby to sleep with her singing/the song”
= lit.: Mary slept the baby (with her singing/the song).

What we observe here suggests, on the extreme take, the possibility that the extent of syntactic malleability of roots is determined by the functional vocabulary available to the native speakers rather than by any inherent limitations imposed by the lexicon. The fact that Slavic languages show more flexibility in this regard than English thus follows from their richer functional structure, in particular, from the presence of AspP, with aspectual morphology ensuring greater malleability of the root.

4. Conclusions

This paper provides an account of the lack of V-to-T movement in East Slavic in which it is argued that while the Russian and Ukrainian data appear to provide an apparently insurmountable challenge to the Rich Agreement Hypothesis, this challenge can be straightforwardly overcome assuming wholesale Roberts' (2010) treatment of head movement as Goal incorporation, coupled with independently needed assumptions about typological differences in the functional sequence and featural specification of heads. The proposed analysis not only removes the ES languages from the list of empirical challenges for the weak RAH, it is straightforwardly extendable to all of Slavic. Finally, I have posited two types of head movement in ES: obligatory V-to-Asp and optional Information Structure-related head raising into the T domain and argued that both are strictly syntactic phenomena. The predictions made by this account, while only briefly discussed here, both further support the account as well as suggest that tracing typological cross-linguistic differences to the presence/absence of Asp in a language's functional sequence is a worthwhile undertaking.

Works Cited

Antonyuk, Svitlana. 2015. "Quantifier Scope and Scope Freezing in Russian." PhD diss., Stony Brook University.

Antonyuk, Svitlana. 2020a. "East Slavic Lexical Aspectual Prefixation as the Locus of Argument Structure Information." Paper presented at XLIX Simposio de la SEL (Sociedad Española de Lingüística), Universitat Rovira i Virgili, Tarragona, Spain.

Antonyuk, Svitlana. 2020b. "The Puzzle of Russian Ditransitives." In Dative Constructions in Romance and Beyond, edited by Anna Pineda and Jaume Mateu, 43–74. Berlin: Language Science Press. DOI:10.5281/zenodo.377653

Antonyuk, Svitlana. 2021. "Russian OVS: Towards a Better Understanding of the Construction's Properties and Significance." Journal of Slavic Linguistics 29: 1–16. FASL extra issue, edited by Andrei Antonenko, John F. Bailyn, Lei Liu, and Anna Melnikova.

Antonyuk-Yudina, Svitlana, and Roksolana Mykhaylyk. 2013. "Prosodic Effects in Word Order: Evidence from Ukrainian." Proceedings of 40th Annual Meeting of the North East Linguistic Society, edited by S. Kan et al., 31–45. GLSA Amherst, MA.

Antonyuk, Svitlana, and Roksolana Mykhaylyk. 2022. "Scope Freezing and Object Shift LQ 8NUDLQLDQ 'RHV ... W\ 0DWWHU"´ Syntax 25:122–146. https://doi.org/10.1111/synt.12229

$UVHQLMHYLü %REDQ. Inner Aspect and Telicity. Utrecht: LOT Publication.

$UVHQLMHYLü %REDQ. ... D ³$ 8QLILHG $QDO\VLV RI 7 ... In Proceedings of ConSole XIV (Vitoria, December 15–17, 2005), edited by Sylvia Blaho, Erik Shoorlemmer, and Luis Vincente, 21–36. Leiden: SOLE.

$ U V H Q L M H Y L ü % R E D Q E ³ 6 O D Y L F K L E H U V U L L [H V $ 197–213. DOI: 10.1163/9789401204446_013.
Bailyn, John Frederick. 1995. “Underlying Phrase Structure and ‘Short’ Verb Movement in Russian.” Journal of Slavic Linguistics 3(1): 13–58.
Bailyn, John Frederick. 2004. “Generalized Inversion.” Natural Language and Linguistic Theory 22: 1–49.
Bailyn, John Frederick. 2005. “Free Word Order and Minimalism.” In St. Petersburg State University Linguistics Papers.
Bailyn, John Frederick. 2012. “The Syntax of Russian.” Cambridge Syntax Guides. Cambridge: Cambridge University Press.
Bobajik, Jonathan David. 2002. “Realizing Germanic Inflection: Why Morphology Does Not Drive Syntax.” The Journal of Comparative Germanic Linguistics 6: 129–167.
% R E D O M L N - R Q D W K D Q ' D Y L G D Q G + | V N X O G X U 7 K U i L Q Better than One.” Syntax 1(1): 37–71.
Borer, H. 2005. Structuring Sense: The Normal Course of Events. Oxford: Oxford University Press.
% R ã N R Y L ü 7 M N R ³ % H F D U H I X O Z K H Q Natural Language D W \ R X and Linguistic Theory 22: 681–742.
% R ã N R Y L ü 7 M N R ³ 2 Q 1 3 Y D Discourse and D X Y H Y .” From Sentence Types to Lexical Categories, edited by G. Grewendorf and T. E. Zimmermann, 179–242. Berlin: Mouton de Gruyter.
Chomsky, Noam. 2001. “Derivation by Phase.” In Ken Hale: A Life in Language, edited by Michael Kenstowicz, 1–52. Cambridge, Mass: MIT Press.
G H Q ' L N N H Q 0 D U F H O D Q G e Y D ' p N i Q \ ³ 5 H W K L Q N Noun Incorporation.” In Syntactic Architecture and Its Consequences: Synchronic and Diachronic Perspectives, Vol. 2, Inside Syntax H G L W H G E \ $ Q G U i V Theresa Biberauer, J. Douglas and Stan Vikner, 465–496. Berlin: Language Science Press.
Dyakonova, Marina. 2009. “A Phase-Based Approach to Russian Free Word Order.” PhD diss., Utrecht University,
Filip, Hana. 1999. “Aspect, Eventuality Types, and Noun Phrase Semantics.” New York: Garland Publishing.
Gribanova, Vera. 2017. “Head Movement and Ellipsis in the Expression of Russian Polarity Focus.” Natural Language and Linguistic Theory 35(4): 1079–1121.
Harizanov, Boris. 2019. “Head Movement to Specifier Positions.” Glossa 4 (1): 140.1–36.
+ D U L] D Q R Y % R U L V D Q G 9 H U D * U L E D Q R Y D Natural ³ : L Language and Linguistic Theory 37(2): 461–522.
Holmberg, Andreas, and Ian Roberts. 2013. “The Syntax–Morphology Relation.” Lingua 130: 111–131.

. D O O H V W L Q R Y D (O H Q D D Q G 5 R X P \ D Q D 6 O D E D N R Y D In Proceedings of FASL 16: The Stony Brook Meeting 2007, 199–214. University of Michigan, Ann Arbor: Michigan Slavic Publications.
Koeneman, Olaf, and Hedde Zeijlstra. 2014. “The Rich Agreement Hypothesis Rehabilitated.”Linguistic Inquiry45 (4): 571–615. DOI: 10.1162/LING_a_00167
Krapova, Iliyana. 1999. “The system of auxiliaries in Bulgarian.”Topics in South Slavic Syntax and Semantics, edited by Mila Dimitrova-Vulchanova and Lars Hellan, 59–90. Amsterdam: John Benjamins. DOI:https://doi.org/10.1075/cilt.172.05kra
. X þ H U R Y i, I. 2007. The Syntax of Gerunds. PhD diss., MIT.
. X þ H U R Y i , Y R Q D 3 * U D P P D W L F D O 1 D W X U D O 1 D Q J X D J H Semantics 20 (1): 1–30.
Lavine, James and Freidin. 2002. “The subject of Defective T(ense) in Slavic.”Journal of Slavic Linguistics10: 253–289.
Matushansky, Ora. 2006. “Head Movement in Linguistic Theory.” Linguistic Inquiry 37:69– 109.
Migdalski, Krzysztof. 2006. The Syntax of Compound Tense in Slavic. Utrecht: LOT Publications, PhD diss., Tilburg University.
Migdalski, Krzysztof. 2013. “Diachronic Source of Two Cliticization Patterns in Slavic.” In Challenging Clitics, edited by Christine Maklenborg Salvesen and Hans-Peter Helland, 135–158. John Benjamins Publishing Company. https://doi.org/10.1075/la.206
Mykhaylyk, Roksolana. 2011. “Middle Object Scrambling.”Journal of Slavic Linguistics 19 (2): 231–72.
Pollock, Jean-Yves. 1989. “Verb Movement, Universal Grammar and the Structure of IP.” Linguistic Inquiry20: 365–424.
4 X D J O L D 6 W H I D Q R % R E D Q $ U V H Q L M H Y L ü 0 D U N R 6 L P review. “Allomorphy, Morphological Operations and the Order of Slavic Verb-Prefixes.” Journal of Slavic Linguistics.
Ramchand, Gillian Catriona. 2008. Verb Meaning and the Lexicon. Cambridge: Cambridge University Press.
Ramchand, Gillian Catriona. 2013. “Argument Structure and Argument Structure-Alternations.” In Cambridge Handbook of Generative Syntax, edited by Marcel den Dikken. Cambridge: Cambridge University Press.
Roberts, Ian. 1993. Verbs and Diachronic Syntax: A Comparative History of English and French. Dordrecht: Kluwer.
Roberts, Ian. 2010. Agreement and Head Movement: Clitics, Incorporation, and Defective Goals. Cambridge, MA: MIT Press.
Rohrbacher, Bernard.1999. Morphology-Driven Syntax. A Theory of V to I Raising and Pro Drop. Amsterdam: John Benjamins.

Slioussar, Natalia. 2007. “Grammar and Information Structure. A Study with Reference to Russian.” PhD diss., Utrecht University.
Slioussar, Natalia. 2011. “Russian and the EPP Requirement in the Tense Domain.” Lingua 121: 2048–2068.
Smith, Carlota. 1991/1997. The Parameter of Aspect. Dordrecht: Kluwer.
Svenonius, Peter. 2004. “Slavic Prefixes inside and outside VP.” Nordlyd 32 (2): 205–253. Special issue on Slavic prefixes, edited by Peter Svenonius. DOI:10.7557/12.68
Tatevosov, Sergei. 2010. “Building Intensive Resultatives.” In Formal Approaches to Slavic Linguistics 18: The Cornell Meeting 2009, edited by Wayles Browne, 289–302. Ann Arbor, MI: Michigan Slavic Publications.
7 R G R U R Y L ü 1 H G D ³ 2 Q W K H 3 U H V H Q F H $ E V H Q F Temporal Interpretation.” PhD diss., University of Connecticut. https://opencommons.uconn.edu/dissertations/1273
Vikner, Sten. 1997. “V-to-I Movement and Inflection for Person in All Tenses.” In The New Comparative Syntax, edited by Vikner and L Haegeman, 189–213. London: Longman.
ä D X F H U 5 R N ³ $ Y S , Q W H U Q D O 5 H V X O W D W L Y H $ Q D Verbal Prefixes.” PhD diss., University of Ottawa.

Deverbal Nominals in Czech: Stem and Root Nominalization

0 L F K D H O D Î D N £ Q \ R Y £

Palacký University, Olomouc, Czech Republic

michaela.cakanyova@upol.cz

Abstract: Czech deverbal nominals have two different kinds of morphological endings. The confusing thing about these is that either form can actually be used as a complex event nominal (CEN) or result nominal (RN). It is the origin of their derivation which can explain this. One type of derived nominals (DN) is derived from the verbal stem and the other type originates directly from the verbal root. These origins result in distinct properties and can account for the more verbal nature of the stem derived CENs and the more nominal nature of the root derived RNs of both morphological types of DNs in Czech. A further analysis also explains how these two types of DNs share some properties and why there can be two morphologically different DNs from the same verb.

Keywords: derived nominals; Czech; roots; stem derivation; feature switch

1. Introduction

There have been many approaches towards deverbal nominals in English and other languages. One of the most influential works on this topic is without doubt Grimshaw (1990). Her analysis of complex event nominals (CEN), event nominals (EN) and result nominals (RN) and their distinct properties is still valid, and many use it as a point of departure for their analyses of various languages. A more generativist perspective is assumed by Borer (1993). In her approach the treatment of DNs is seen as two different kinds of derivation, one of which is lexical, meaning that the noun is derived at the level of lexicon before being inserted in the D structure, and the other one is syntactic, when the noun can be derived later after the insertion into D structure. She applies this to English CENs, which she calls process nominals, and claims that these are derived post syntactically while English RNs are derived pre-syntactically. The lexical pre-syntactic

derivation puts RNs on the par with non-derived nominals and could help explain why RNs have purely nominal properties. CENs are derived through a null verbal head which moves to the N head affix.

In her subsequent work Borer (2003) turns her focus on the presence or absence of the argument structure (AS) and distinguishes between AS nominals and referential nouns. This means that only the nouns derived syntactically from verbs can have the argument structure from the verb and the suffix itself does not carry any verbal features. The derivation is still taking place at two different levels – for this reason she uses the term parallel morphology. This approach differs from the Distributed Morphology (DM) framework namely in the pre-syntactic level of insertion, which is something not acceptable for DM. In the DM framework Marantz (2000, 2001) treats the derivations as originating from the roots with the application of abstract categorizer, *v* or *n* to yield the particular part of speech. These categorizers are phonologically empty.

Borer (2014) updates her analysis and joins the two ideas. To distinguish between AS nominals and R nominals she deems it necessary to consider, on the one hand, the level of derivation and on the other hand the combination of roots with categorial functors (C-functors). Both AS and R nominals (those without AS) are morpho-phonologically identical. The difference between the two lies according to Borer in their "distinct syntactic properties of the larger nominal constituent within which the derived nominal is embedded" (2014, 80). She applies a top-down, constructionist approach to event/argument structure because she claims that the argument structure cannot originate from the embedded verb nor can it emerge from the root. Roots are devoid of any semantic and syntactic properties; they serve as mere phonological indices and moreover they do not have arguments.[1] The derivation of R nominals happens when a root gets embedded within a C-structure with a particular C-functor, which results in their category. The structure of AS nominals is more complex with up to two additional layers, X and Y of the extended verbal projection. They layer Y is optional, licensing the internal argument, and the layer X licensing an event argument or even external argument.

The most important part of her study concentrates on the compositionality of the DNs. AS nominals are seen as always compositional, i.e., they include a verbal stem and also the nominalizing suffix. R nominals, on the other hand, can be either compositional or non-compositional. This is not surprising since "[e]vent denotation, as such, is not restricted to AS-nominals and is rather found in underived nominals as well" (Borer 2014, 72). An example of that would be nouns such as class or wedding (EN according to Grimshaw) which can have the eventive meaning. Their duration cannot be expressed through a PP (*the class for two hours), but they need to follow that verb, e.g., the class lasts two hours.

1 This is in contrast with Harley (2009), who claims that roots can actually take internal arguments.

In our analysis, we are going to build on Borer's and DM analyses, but we will need to adjust these according to the particularities of Czech DNs.

2. Czech Deverbal Nominals

The Czech deverbal nominals are traditionally divided into two categories based solely on their final suffix. The two types of derived nominals in Czech according to their endings are:

- type 1: has the ending -ní/-tí as in balení 'packaging' or psaní 'writing'
- type 2: has a greater variety of possible nominalizing endings (including zero suffix), but the most typical suffixes include -ba/-ka as in malba 'painting' or þ H W E D 'reading'

They are typically further specified as to their semantics. Type 1 nouns tend to refer to events and actions. Type 2 nouns refer either to agents, instruments, results but also states and events.

(1) X þ H Q t P \ W t Y \ ã t Y i Q t] N R X ã H Q t
'teaching, washing, embroidering, examining'

(2) X þ L W H O P \ þ N D Y ê ã L Y N D] N R X ã N D
'teacher, dishwasher, embroidery, examination'

, Q W K H S U H Y L R X V O L W H U D W X U H 3 D Q H Y R Y i . D U O have been thoroughly treated mostly with regard to their dis(similar) morphosyntactic properties, but there has been a deeper syntactic analysis lacking. Namely, their derivation processes were not paid sufficient attention to and the explanation of their different syntactic properties thus fell short. In this paper I am trying to address this gap from the point of view of their different origin, and in the light of that I am trying to explain their syntactic behavior which is sometimes different and sometimes almost identical.

These two types of nominals have many similarities with most regular non-derived Q R X Q V) R O O R Z L Q J . D U O t N D Q G 1 • E O H U D Q G . D U O include: the ability to appear after a preposition (3), the ability to follow a determiner (4) and allowing for a relative clause post-modification (5)–(6).

(3) E H]] N R X ã N \ ] N R X ã H Q t
'without an exam/examination'

(4) W D] N R X ã N D W R] N R X ã H Q t
'this exam/examination'

(5)]NRXãND NWHUi WUYDOD KRGLQX
‘an exam which lasted an hour’

(6)]NRXãHQt NWHUp WUYDOR KRGLQX
‘examination which lasted an hour’

Regarding their inner structure, both type 1 and type 2 can (C)ENs with some parts of the AS and RNs. These two kinds of nominals differ in their more or less verbal or nominal properties. The RNs have strictly nominal properties, which means that they resist aspect (7), resist instrumental (INS) agent (8), can be modified by genitive (GEN) and a PP (9), are countable (10), and do not require an internal argument (11):

(7) (napsat) QDSVDQt Y\VEtUDW Y\VEtUND
write. ‰ ~ ‹ • writing ‰ ~ ‹ • collect. ‰ ~ ‹ • collection. ‰ ~ ‹ •

(8) SVDQt 3HWUHP VEtUND 3HWUHP
writing Petr. , ‡ Œ collection Petr. , ‡ Œ

(9) SVDQt PRMt VHVWU\VEtUND QDãHKR NRVWHOD
writing my. € ~ sister. € ~ ‡ collection our. € ~ church. € ~ ‡
‘a letter of my sister’ ‘a collection of our church’

(10) GY SVDQt W L VEtUN\
‘two writings’ ‘three collections’

(11) (a) 3HWU VEtUi PLQFH 3HWU SVDO -DQ
Petr collects coins.acc Petr wrote Jana.dat
‘Peter collects coins’ ‘Peter wrote to Jane’

(b) 3HWURYD VEtUND 3HWURY SVDQt -DQ
Petr’s collection (coins.gen) Petr’s writing (Jana.gen)
‘Peter’s collection of coins’ ‘Peter’s letter to Jane’

(C)ENs on the other hand can keep the original verbal argument structure and are therefore less nominal than RNs. There are many examples of (C)ENs of type 1 and 2 from the

2 With type 2 I am not using CEN but only EN as the type of event depicted by this kind of QRPLQDO LV VOLJKWO\ GLuHUHQW WKDQ WKH &(1 QDPHO\ IXUWKHU GLuHUHQFHV EHWZHHQ &(1V DQG (1V LQ VHFWLRQ derivation process.

same verb used in similar contexts (12)–(13). However, their structure is not identical. If type 2 expresses an EN, then it can keep the argument structure of the verb – the object is realized as an NP in GEN.[3] For type 1 this is necessary (with transitive verbs) (13).

(12) 3 U R G i Q t bytu zabralo Petrovi F H O po odpoledne.
selling.| ~ ‡ • ' ‰ ~ À D W € took Peter.} z • whole afternoon
µ 6 H O O L Q J W K H À D W W R R N 3 H W H U Z K R O H D I W H U Q R

(13) Prodej (bytu) zabral Petrovi F H O po odpoledne.
sale. ~ ‡ • ' ‰ ~ À D W € ~ took Peter.} z • whole afternoon
µ 7 K H V D O H R I W K H À D W W R R N 3 H W H U Z K R O H D I W

This is but one of many differences between the types 1 and 2 (C)ENs. This and further differences which will be presented all stem from the nature of the type of derivation of these nominals.

Let us take a look at type 1 CENs first. These nominals can be derived from almost any verb (except for modals and some stative verbs). They form a bigger category including mass nouns in singular neuter only. It is a highly productive category including nominals derived from loan verbs. The loan verbs are easily transformed into nominals just by adding the suffix -ní as in the case of verb sharovat 'to share' in infinitive and the DN sharování 'sharing'. Type 2 ENs typically do not allow for this as their derivation is not that productive. They come in different kinds of genders as they have full grammatical paradigms.

3. Two Different Derivations of Czech DNs

The Czech DNs of both morphology types can and do function as (C)ENs and RNs, but the above-mentioned analysis suitable for English does not seem to work too well for Czech. The problem is that type 1 visibly keeps a part of the verbal stem while type 2 does not seem to have enough verbal properties to be deemed derived from a verb at all, but instead it seems that it is derived directly from the root. So, the question is, how it is possible that one form can function as both RNs and (C)ENs.

Deverbal nominals in Czech originate either from (verbal) stem derivation or directly from the root. In case of type 1 (ní/tí) we talk about the stem derivation as in: V_{stem} : & (1 : 5 1. This can help account for the fact that all type 1 DNs are primarily CENs and only some of these can function as RNs as well. In the case of type 2 (ba/ka) we are talking about a categorial ROOTs in the sense of Halle and Marantz (1993). The root combines with a nominalizing n head and becomes a RN. The next possible step

3 I am using a NP rather than a DP for Czech because unlike English, Czech does not require the determiner layer for singular nouns, the D layer is optional.

is the eventive use of this DN: : 5 1 : (1 This results in the fact that all type 2 DNs are first and foremost RNs and only some of these can function as ENs.

The stem derivation of type 1 is a sequential procedure. Following Borer (2014) I claim that first the CEN is derived from the verbal stem and then in some cases a AS is derived from the CEN. This kind of consecutive derivation explains why there is no RN of type 1 that would lack its CEN type 1 counterpart. The stem-derivation of CEN K D S S H Q V D W W K H V \ Q W D F W L F O H Y H O S K D V H I L Q D O R

(14) vybavit Y \ E D Y H Q t Y \ E D Y H Q t
equip. , ‡ • equipping.| ~ ‡ equipment‹ ‡

The thematic affix present in these DNs is what distinguishes them from the other type of ENs and is responsible for their more verbal properties.

The most difficult question is how one and the same form can function as two different kinds of nominals. I claim that the initial derivation process is the same for CENs and RNs of type 1; however, the CEN can be 'coerced' (Harley 2009) into becoming a RN.[4] So, the next step involves a kind of a 'feature switch' through the uninterpretable features 3 D Q D J L R W L G L V D Q G * U R K P D Q Q + D Y U D Q R Y i R I morpheme added to the CEN but it simply switches its category.[5]

(15) å H K O Lå W K O H Q Ø
iron. , ‡ • ironing. | ~ ‡

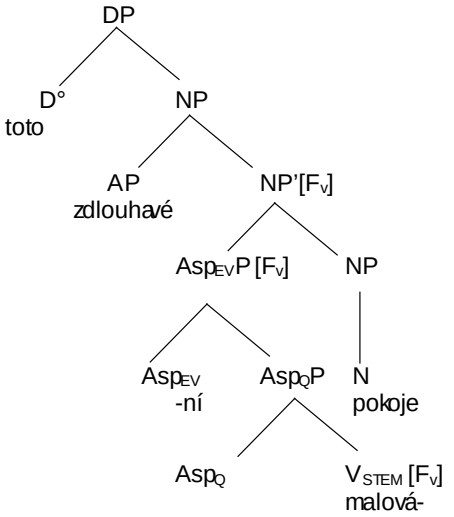
CEN of type 1 in a tree schema which shows the stem derivation and the retained aspectual structure.

(16) toto] G O R X k D Y á ní pokoje
'this lengthy painting of the room'

---

4 Harley (2009, 338) interestingly notices that just as mass nominals can frequently become countable nominals, so can AS nominals get the resultative reading, acquiring the capacity to express grammatical number. This process (and the Num head) as a by-product excludes the presence of a syntactic object.

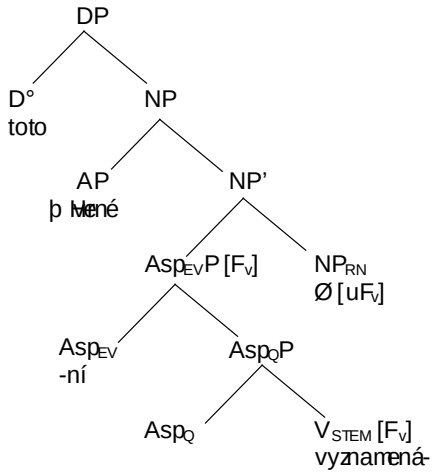
5 There have been prior attempts at the explanation of the seeming polysemy of DNs in various languages. Bierwisch (1990) comes up with the idea of the conceptual shift applicable to German '1 V Z K L F K R U W D R P D Q W L F O H [L F D O G U L I W R I F H U W D L Q L W H P V ± categorial switch idea. The applicability of the shift depends on the semantic-syntactic properties of the given verb. Another such example is Pustejovsky (1995), who treats DNs as polysemous, to be precise, complementary polysemous. This is a property that other non-derived nouns have as well. We can see this in the case of zero morpheme derivations of e.g., to call and a call or in metaphorical use of nouns.



) L J X U H Czech CEN.

RN of type 1 in a tree schema including the categorial feature switch from CEN to RN.

(17) W R W R ‘this red award’



) L J X U H Czech RN of type 1.

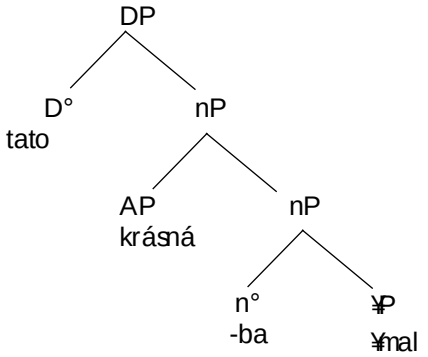
The other type of the nominalizing process is direct root nominalization yielding a type 2 RN. The process of type 2 nominalization must be different from that of type 1 as not all RNs of type 2 have their EN counterpart. The process is a nominalization in the sense of Marantz (2001) and Harley (2008) where the acategorial root merges with the specific type of nominalizing suffix n°. And also here, in order to get a corresponding EN, the feature switching is applied.

(18) malovat ¥ mal-ba malba
paint., ‡ • picture.‹ ‡ painting.~ ‡

(19) V N R þ L ¥W sk-ok Ø
jump. , ‡ • jump. ‹ ‡

RN of type in a tree schema showing the nominalization from the root.

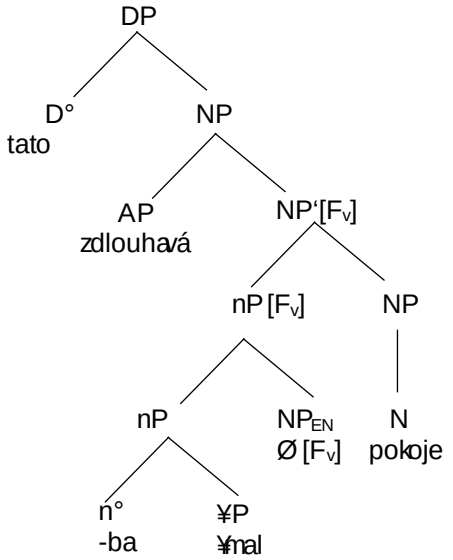
(20) W D W R malba i V Q i
‘this beautiful painting’



) L J X U H Czech RN of type 2.

EN of type 2 in a tree schema showing the categorial switch from RN.

(21) W D W R] malba þ kOY i
‘this lengthy painting of the room’



) L J X U H Czech EN.

The properties of both morphological types of Czech DNs can be summarized in the following table:

	6 X v [	& (1	(1	5 1	Derivation
Type 1	ní/tí	all type 1	—	some type 1	6 W H P 9 : & (1 : 5 1
Type 2	ba/ka	—	some type 2	all type 2	5 R R W 5 1¥::(1

7 D E O H Summarizing table

4. Different Properties Based on the Particular Kind of Derivation

The RNs of either type of DNs seem to share rather similar purely nominal properties. However, depending on the origin of the (C)EN, we can see their different behavior. This is specifically applicable to their argument structure. Based on the analysis above, it is not surprising to see that type 1 CENs retain more verbal features than type 2 ENs; these features namely include the ability to take reflexive pronouns (22), to express repetition and aspect (23), and take negation (24), and the ability to be modified by As expressing frequency (25).

&(1 W\SH YV 7\SH

7\SH PDORYiQt VH µSDLQWLQJ RQHVHOI¶
Type 2: *malba se

7\SH VEtUiQt KRGLQX]D KRGLQX
‘picking for an hour’
Y\VEtUiQt KRGLQX]D KRGLQX
‘having picked in an hour’

Type 2: stavba (hodinu / za hodinu)
(innert) ‘building for / in an hour’
dostavba (hodinu / za hodinu)
‘having built for / in an hour’

7\SH QHPDORYiQt QHVEtUiQt
‘not-painting, not-picking’
7\SH QHPDOED QHVE U

7\SH þDVWp PDORYiQt
‘frequent painting’
7\SH "þDVWi PDOED
‘frequent painting’

Moreover, only the CENs, which denote an event, can keep the argument structure, which corresponds to the argument structure of the verb, i.e., they have both the internal and external argument. The ENs of type 2 can keep just the internal argument, but they do not necessarily need it (27).

(26) 3 HGQiãHQt OiWN\ QDãHPX profesorem).
presenting| ~ ‡ topic. € ~ ‡ our. , ‡ Œ professor. , ‡ Œ
‘Presenting the topic by our teacher’

(27) (a) 'QHãQtmalba]iWLãt *(studenty) trvala hodinu.
today’s painting.~ ‡ still-life. € ~ ‡ students. , ‡ Œ took hour
‘Today’s painting of the still life (by the students) took an hour.’

(b) 'QHãQtmalba ___ trvala hodinu.
today’s painting.~ ‡ took hour
‘Today’s painting ___ took an hour’

Perfective verbs (PERF) and nouns alike need objects to be expressed (28) unless they are inherently understood, while imperfective (IMPF) verbs and nominals derived from them do not need them but can have them included (29). Since the perfective aspect is not compatible with type 2 ENs (30), this only applies to type 1 CENs.

(28) 3 H P D O R Y iSQR P D O R*(domu) Q i P zabralo F H O podpoledne.
Repainting.‰ ~ painting.‰ ~ ‹house us took all afternoon
'Repaintintg/painting the house took us all afternoon.'

(29) 0 D O R Y i Q t(domu) Q i P zabralo F H O podpoledne.
painting., † ‰ house us took all afternoon
'Painting the house took us all afternoon.'

(30) *Pomalba (domu) Q i P zabrala F H O p odpoledne.
Painting.~ ‡ ‰ ~house us took all afternoon

According to Grimshaw (1990), only CENs could be modified by modifiers which indicate agenthood as the external argument is somewhat preserved in the structure. The ENs do not seem to operate too well with such modifiers.

(31)] i P U Q p S R U X ã R Y pirQvidel
intentional violating.| ~ ‡ rules. € ~ ‡
'an intentional violation of rules'

(32)] i P U Q i porucha W H F K Q L F]NDp K]RQ t
intentional failure technical.€ ~ ‡ mechanism€ ~ ‡

Another of Grimshaw's original tests includes the applicability of adverbial modifiers. If we apply such modifiers to Czech CENs and ENs, we can see the difference at first glance. Even though it is mostly more natural to use adjectival modifiers for both types of DNs, it is downright ungrammatical to modify ENs with adverbials.

(33) (a) R G S U H] H Q W R Y i Q t W p P D W X M D V Q D V U R] X P L W H O
'presenting the topic clearly and understandably'

(b) M D V Q p D V U R] X P L W H O Q p R G S U H] H Q W R Y i Q t W p P D
'clear and understandable presenting of the topic'

(34) (a) M D V Q i D V U R] X P L W H O Q i S U H] H Q W D F H W p P D W X
‘clear and understandable presentation of the topic’

(b) S U H] H Q W D F H W p P D W X M D V Q D V U R] X P L W H O Q
presentation (topic€ ~)‡ clearly and understandably

5. The Arguments of Czech CENs

Depending on the verb’s valency the DN can have zero, one or two arguments. DNs derived from mono-transitive verbs have the patient in genitive case and the agent is typically expressed by an instrumental case NP. Both arguments appear post nominally, and the agent assumes the position of an adjunct.

(35) (a) S R V H N i Q tzahrady] D K U D G Q t N H P
mowing garden. € ~ ‡ gardener. , ‡ Œ
‘gardener’s mowing the garden’

(b) prodej domu *otcem
sale house. € ~ ‡ father. , ‡ Œ

Moreover, if the agent is animate and singular, it can appear pre-nominally as a possessive NP. This works for both CENs and ENs and also any other non-derived noun.

(36) (a)] D K U D G Q t SRVHNzahrady
gardener’s mowing garden. € ~ ‡
‘gardener’s mowing the garden’

(b) R W F $ Y prodej domu
father’s sale house. € ~ ‡
‘father’s sale of the house’

If there is only one argument present, then it is in genitive case and the meaning can be ambiguous as to whether it is the agent or the patient, especially if both NPs are animate.

(37) S R N R X ã H Q t
temtping Jane. € ~ ‡
‘Jane’s tempting/Jane is being tempted’

It is, however, necessary to add that the patient (verbal object) is not compulsory for imperfective DNs with argument structure, only for the perfective ones as we have seen L Q í

With ditransitive verbs in Czech, we expect to get two NPs, one in the accusative and one in dative case. These verbal objects typically fulfill the roles of theme/patient and recipient respectively. After nominalization, the cases change from accusative to genitive for the patient and dative remains the same for the recipient. This is not surprising as it is the structural case which changes while the (inherent) lexical case stays the same. There is also a left-right order of cases in Czech: NOM-ACC-GEN-DAT-INS (cf. Caha 2009, 47). So, the GEN needs to precede the DAT. The theme/patient is the obligatory argument of the verb and in the nominalized structure it is an obligatory GEN case marked argument of the DN.

(38) 6 W D U R V W D O H W R V G D U R Y D O G W H P K U D þ N \
'The mayor gave (children) toys this year.'

(39) ' D U R Y i Q t K U D þ H N G W H P S L Q H V O R V W D U R V W R
'Giving the toys (to children) secured the mayor new votes.'

(40) Eva napsala (Adamovi) dopis.
'Eva wrote a letter (to Adam).'

(41) 1 D S V i Q t dopisu $ G D P R Y L W U Y D O R (Y W ê G H Q
'Writing the letter (to Adam) took Eva a week.'

The CENs with AS (unlike ENs) are compatible with different aspects through prefixes but also through the stem *vá–* morpheme, which triggers the imperfective aspect.

(42) (a) zastavit zastavovat
stop. ‰ ~ ‹ • stop. , † ‰ •

(b)] D V W D Y H Q t] D V W D Y R Y i Q t
stopping. ‰ ~ ‹ • stopping. , † ‰ •

The aspect layer of the CENs has two consequences, firstly that nouns derived from transitive perfective stems have to have overt patient expressed as an NP (28)–(29). This holds in almost all cases with the exception of understood objects with a limited number of verbs, such as vytírání 'mopping' or å H K O H Q t 'ironing' as it is assumed that the unexpressed objects are vytírání podlahy 'mopping the floor' and å H K O H Q t S U i G O 'ironing the laundry'.

$ F F R U G L Q J W R ' Y R i N W K H &] H F K S ' 1 V R L Y W \ S H L F L Z O W (n/t can actually be found in Czech passives) and are thus unable to check the accusative case or case mark anything at all. That is the reason for the genitive case.

6HFRQGO\ 'YR iN QRWLFHV WKDW WKH S
inability to combine with durative adverbials. We have already seen that type 1 CENs can combine with time adverbials such as in an hour / for an hour. But this applies only as long as they are derived from an imperfective verb (43). When a CEN is derived from a perfective verb it cannot combine with durative adverbials but only with the terminative ones (44).

(43) P DO RY iQ t MHGQRKR REUD]X]D KRGLQX KRGLQX
'painting one picture in an hour / for an hour'

(44) Q DP DO RY iQ t MHGQRKR REUD]X]D KRGLQX KRGLQ
'having painted one picture in an hour'

The DNs derived from intransitive verbs also mirror their inner structure, unergative verbs like arrive or stutter have just the external argument, which is the agent.

(45) Jan koktal
Jan. ‡ ˆ † stuttered
'Jan stuttered'

(46) N R N W i QÍ Jana/*Janem
stuttering Jan. € ‚/*Jan. , ‡ Œ
'Jan's stuttering'

Another subclass of intransitive verbs are unaccusatives such as to fall. These do not express the agent in the subject position but rather a patient or theme.

(47) G G H þ ‰ ‹ Nl
'a grandfather fell'

(48) X S D G QÍ X WÍ H þ N D G G H þ N H P
falling grandfather. € ‚/grandfather. , ‡ Œ
'grandfather's fall'

These DNs mirror the argument structure of the verbs they are derived from, namely not being compatible with the agents expressed in an INS case phrase.

6. Conclusions

The DNs are of three different syntactic types in Czech, complex event nominals, event nominals and result nominals. RNs of both suffix type 1 and 2 show identical (nominal)

properties and they do not retain any verbal features. CENs of type 1 suffix retain more verbal properties because they are derived from the verbal stem and ENs of type 2 are less verbal because they are derived through RNs. This means that they do not have the verbal argument structure. They are at best eventive nominals on the par with some eventive non-derived nouns such as hodina ‘class’ or svatba ‘wedding’. The verbal stem derivation of CENs and the root origin of the RNs explains their different syntactic properties.

References

Bierwisch, Manfred. 1990. “Event Nominalizations: Proposals and Problems.” Acta Linguistica Hungarica í ±

Borer, Hagit. 1993. “Derived Nominals.” Ms. University of Massachusetts, Amherst.

Borer, Hagit. 2003. “Exo-skeletal vs. Endo-skeletal Explanations.” In The Nature of Explanations in Linguistic Theory, edited by J. Moore and M. Polinsky, 31–67. Chicago: CSLI and University of Chicago Press.

Borer, Hagit. 2014. “Derived Nominals and the Domain of Content.” Lingua 141: 71–96.

Caha, Pavel, 2009. “The Nanosyntax of Case.” PhD Diss. University of Tromsø.

'YRiN 9UD ³&DVH $VVLJQPHQW $VSHFW DQG In Noun Valency HGLWHG E\ 2OJD 6SHYDN í $PVWHU

Emonds, Joseph. 2000. Lexicon and Grammar: The English Syntacticon. Mouton de Gruyter: New York.

Grimshaw, Jane. 1990. Argument Structure. Cambridge, MA: MIT Press.

+DOOH 0RUULV DQG $OHF 0DUDQW] ³'LVWULEXWH tion.” In The View from Building 20 HGLWHG E\ . +DOH DQG 6 - Cambridge, MA: MIT Press.

Harley, Heidi. 2008. “Compounding in Distributed Morphology.” In Oxford Handbook of Compounding, edited by R. Lieber and P. Stekauer, 129–144. Oxford: Oxford University Press.

Harley, Heidi. 2009. “The Morphology of Nominalizations and the Syntax of vP.” In QP and Nominalizations, edited by M. Rathert and A. Giannakidou, 320–342. Oxford: Oxford University Press.

+DYUDQRYi . DW “The Long Nominalization Process in English, Dutch, and Czech.” PhD diss., Palacký University.

.DUOtN 3HWU ³9DOHQFH VXEVWDQWLY Y PRGL í XQLYHU]iOLD D VSHFLILND (GLWHG E\ = +DG Masarykova univerzita.

.DUOtN 3HWU ³-HãW MHGQRX NþHýHNiWVNpKR – univerzália a specifika HGLWHG E\ =GHQD +DGNi DQG 3HW /LGRYp QRYLQ\

.DUOtN 3HWU DQG 1RUEHUW 1•EOHO Slovo a slovesnost 59:105–112.

Marantz, Alec. 2000. “Roots: The Universality of Root and Pattern Morphology.” Paper presented at the Conference on Afro-Asiatic Languages, University of Paris VII, June 2000.

Marantz, Alec. 2001. “Words.” Paper presented at the West Coast Conference on Formal Linguistics, University of Southern California Los Angeles, 24 February 2001. Available online at https://babel.ucsc.edu/~hank/mrg.readings/Marantz_words.pdf

3 D Q D J L R W L G L V 3 K R H Y R V D Q G . O H D Q W K H V . - * U R K P D rial Switches and Prolific Domains. Linguistic Analysis 35: 141–161.

3 D Q H Y R Y i - D U P L O D ³ 3 R] Q i P N \ N ý Y H G ô W H Q F L ± S R G V W zália aspecifika 2 H G L W H G E \ = G H Q D + O D G N i D Q G 3 H W Masarykova univerzita.

Pustejovsky, James. 1995. The Generative Lexicon. Cambridge, MA: MIT Press.

9 H V H O R Y V N i / X G P L O D ³ . D Q D O ê] H þ H ý V H Q V W Y R Q D H Y í univerzália aspecifika H G L W H G E \ = G H N D + O D G N i D Q G 3 H Masarykova univerzita.

Allomorphy Conditioned by Post-Linearization Spanning: Evidence from Italian Theme Elements

Pamela Goryczka

University of Vienna, Vienna, Austria

pamela.goryczka@univie.ac.at

Abstract: This contribution provides empirical support from Italian for span-conditioned theme (vowel) allomorphy. Crucially, following Merchant's (2015) Span Adjacency Hypothesis and Haugen and Siddiqi's (2016) proposal of adopting post-linearization spanning for non-lexicalist realizational models of morphology such as Distributed Morphology (DM, Halle and Marantz 1993), it is argued that the preferable model of spanning is indeed the one which targets linearly adjacent formal features without reference to structural constituency. With regard to previous analyses involving Italian theme formatives (Embick 2016, Calabrese 2019), the analysis presented here manages to hold without having to resort to additional morphological operations such as fusion and impoverishment.

Keywords: allomorphy; theme formatives; distributed morphology; spanning

1. Introduction

In recent years, research on the Romance verbal system has often focused on providing evidence for the presence of paradigmatic patterns known as morphomes (Maiden 2004, 2005, 2018). From this perspective, Italian third conjugation verbs displaying the so-called -isc- augment are considered to represent the type of distributional regularity expressed by one of the most prominent morphomes in Romance verb morphology known as the N-pattern (Meul 2010, 2013, Da Tos 2013).[1] The augment appears only in

1 According to Maiden (2009), the source of the N-pattern is in fact phonological, i.e., brought on E\ WKH TXDOLW\ GLuHUHQWLDWLRQ EHWZHHQ VWUHVVHG DQ morire 'die' wherein the Present Tense singular forms, the 3 ‰ … and the 2 Œ € imperative share a roo distinct from the rest of the paradigm (e.g., 1 Œ € 'I die' vs. 1 ‰ 'we die'). Apparently the pattern then appealed also to other verbs whose root alternation was not due to phonological reasons but which nevertheless showed the same distributional pattern.

the Present Indicative and Subjunctive forms – except for the 1 ‰ … and the 2 ‰ … – as w
as the 2 Œ € Imperative (cf. Table 1).

	Present Indicative	Present Subjunctive	Imperative
1 Œ	fin-[isk]-o	fin-[isk]-a	
2 Œ €	I L Q > L U U @ L	fin-[isk]-a	fin-[isk]-a
3 Œ €	I L Q > L U U @ H	fin-[isk]-a	
1 ‰	fin-[ja]-mo	fin-[ja]-mo	
2 ‰	fin-[i]-te	fin-[ja]-te	fin-[i]-te
3 ‰	fin-[isk]-o-no	fin-[isk]-a-no	

Table 1. Distribution of the augment in Italian*finire* ‘end’[3]

However, while this observation may hold true for standard Italian, some Italo-Romance varieties, such as Ligurian, can give us a more varied and contrasting picture of the intra-paradigmatic distribution of -isc-, thus dispersing the original pattern of distribution: L Q 7 D E O H W K H Y D U L D Q W L U L K D V E H H Q H [W H Q
Present Subjunctive. The same pattern can also be observed in many Lombard dialects (e.g., Ticinese), in Upper Engadinian and is additionally found in Corsican and in related dialects in northern Sardinia (Meul 2010, 14).

	Present Indicative	Present Subjunctive
1 Œ €	S D W > L U L @ X	S D W > L U @ H
2 Œ €	S D W > L U @ L	S D W > L U @ L
3 Œ €	S D W > L U @ H	S D W > L U @ H
1 ‰	pat-[i]-mu	S D W > L U L @ P X
2 ‰	pat-[i]-i	S D W > L U @ L
3 ‰	S D W > L U @ H	S D W > L U @ H

Table 2. Ligurian*patí* ‘suffer’ (Meul 2010, 14)

2 For reasons of space, this paper focuses on the augment’s presence in the Present Tense. As for the Imperative, a more in-depth discussion remains a desideratum of further investigation.

3 The velar /k/ in /isk/ is palatalized when it occurs before front vowels, therefore yielding the Y D U L D Q W L U R U P R U H S U H F L V H O \ Œ € and in the Œ € L Q W H U Y R F D

Finally, in some dialects such as Milanese the -isc- augment has also extended to the Conditional and Future tenses. As can be observed in Table 3, the elements /isa/ and /i/ seem to be paradigmatically complementary – either cap-i-róo or cap-iss-a-róo ‘I will understand’ is being used – but syntagmatically incompatible (*cap-iss-i-roo). The sequences /isa/ and /i/ seem thus mutually interchangeable which prompts Meul [4] WR FRQFOXGH ³$V IRU LWV µVWDWXV¶ WKH L morph of the thematic vowel /i/ of the fourth conjugation.”

	Future	Conditional
1 Œ	cap-i-róo / cap-iss-a-róo	cap-i-ría / cap-iss-a-ría
2 Œ	cap-i-rée /cap-iss-a-rée	cap-i-ríet / cap-iss-a-ríet
3 Œ	cap-i-rá / cap-iss-a-rá	cap-i-ría / cap-iss-a-ría
1 ‰	cap-i-rémm / cap-iss-a-rémm	cap-i-ríom / cap-iss-a-ríom
2 ‰	cap-i-ríi / cap-iss-a-ríi	cap-i-ríov / cap-iss-a-ríov
3 ‰	cap-i-ránn / cap-iss-a-ránn	cap-i-ríen / cap-iss-a-ríen

Table 3. Milanese capí ‘understand’

Crucially, Meul’s concluding remark refers exclusively to the Milanese data. In what follows, I will address the possibility that the -isc- augment does indeed function as a theme allomorph, however not only in some Italo-Romance varieties but also in standard Italian.

2. Theme Vowels in Italian

Traditional analyses of Romance verbs typically divide a verb form into root, theme vowel, and inflectional suffixes. The theme vowel (or thematic extension) determines the morphophonological surface shape of the verb, however it does not contribute to its syntactic or semantic meaning. The theme vowel thus seems to facilitate the recognition of class membership as a “purely morphological marker” (Oltra-Massuet 2020). Across frameworks it is described in similar ways, as “morphological glue” (Anderson 1992), “empty morph” (Aronoff 1994) or “ornamental piece” (Embick 2010).

4 In some varieties of Lombard (as in the Milanese example above), the presence of the augment JRHV WRJHWKHU ZLWK WKH JHQHUDOL]DWLRQ RI WKH ¿UVW the form /isa/ (cf. Meul 2010).

Based on the form of the infinitive, Italian verbs are usually divided into three conjugation classes.[5] Verbs ending in -are (e.g., cant-a-re ‘sing’) constitute by far the largest class. Moreover, the first conjugation class is highly productive since new verb creations are more likely to belong to this class than to any other (e.g., chattare ‘chat’, instagrammare ‘instagram’, selfarsi ‘take a selfie’). It also contains mainly regular verbs, i.e., verbs that have only one single stem for all tense forms. On the other end of the spectrum, second conjugation class verbs in -ere (prend-e-re ‘take’, tem-e-re ‘fear’) account for the smallest group; they are mostly comprised of less frequent and irregular verbs. Finally, third conjugation verbs in -ire can generally be divided in verbs presenting the augment -isc- in the Present Tense forms (and the 2 Œ € imperative) except for the 1 ‰ … and 2 ‰ … (fin-isc-o ‘I finish’, cf. Table 1) and (a much smaller group of) verbs that are formed without the augment (e.g., dorm-i-re ‘sleep’).[6] Although assumed to be little productive as a whole, new (mostly parasynthetic) verb formations presenting the -isc- augment have been introduced to this class over the years (e.g., ingiallire ‘turn yellow’, abbellire ‘improve, adorn’).

The theme vowel can easily be identified in the infinitive since it is located between the root and the inflectional suffixes. In other forms, however, the theme vowel given of a conjugation class may be missing (e.g., Present Tense 1 Œ € ‘sing’, prend-o ‘I take’, dorm-o ‘I sleep’)[7] or may present an allomorph of the original theme vowel (e.g., Present Tense 3 ‰ … cant-a-no ‘they sing’ vs. prend-o-no ‘they take’, dorm-o-no ‘they sleep’). Comparing any non-augmented verb with an augmented verb, it becomes rapidly clear that the presence of -isc- actually facilitates the classification of a verb. The surface forms of 1 Œ € canto, prendo, temo and dormo may as well belong to any of the conjugation classes, but finisco immediately evokes recognition of third conjugation affiliation.

It thus seems that -isc- functions as a very strong identity marker. Just as the theme vowel, the augment -isc- is a morphological formative necessary in some forms of the paradigm to form a stem before adding inflectional suffixes. Just as the theme vowel, it does not have any syntactic or semantic properties that could impinge on the overall

---

5 There are alternative proposals such as the division of the verbs into two macroclasses: the ¿ U V W F O D V V F R Q W D L Q H V H U H V G L Y L G H G L Q V H F R Q G D U L H U H (Dressler and Thornton 1991). However, many other similar analyses usually divide the second macroclass in more subgroups (Vincent 1988, Schwarze 1999). For the purpose of this paper, the exact number or type of conjugation subclass is not really relevant; in any case, all proposals separate, at some point, -ere verbs with augment from -ire verbs without augment.

6 In a recent study, Da Tos (2013, 47) analyzed 450 (non derivated) verbs in -ire and concluded that only 19 were in fact non-augmented verbs.

7 According to Scalise (1994, 155), the theme vowel is present in the underlying form but due to an ubiquitous vowel deletion rule the theme vowel is deleted when followed by another vowel (e.g., cant-a-o > cant-o ‘I sing’).

meaning of the verb. Just as the theme vowel, it seems to be a morphological marker that instantiates conjugation class membership. It thus seems plausible to hypothesize that -isc- functions as a theme and is therefore located in the theme position of the verb.

But one last minor issue still needs to be addressed before the hypothesis can be implemented in a formal analysis. The vowel /i/ in the sequence -isc- and the third conjugation theme vowel /i/ are homophonous. It is important to stress that the /i/ vowel in -isc- is not the theme vowel /i/. This can easily be shown by looking at the 3 ‰ …: in the augmentless form-o-no the theme vowel allomorph /o/ is present (just like in second conjugation verbs such as prend-o-no); in the augmented fin-isc-o-no, the same allomorph /o/ is also present. Since /isk/ is not situated in the verbalizing head (little v), only remains one position and that is the theme position. To sum up, the theme elements we are actually dealing with in third conjugation verbs presenting the augment are /iski/ DQG LUUL ZLWK WKH VHFRQG YRZHO L JHWWLQJ HLV (e.g., 1 Œ € /fin-isk(i)o/, 2 Œ€Q LUU L L RU XQGHUJRLQJ D SURF fin-isko-no/) or lowering (e.g, 3Œ € ILQ).LUUH

3. Theoretical Background

I propose an analysis of Italian third conjugation verbs presenting the augment -isc- within the theoretical framework of Distributed Morphology (DM), as introduced in Halle and Marantz (1993, 1994) and modified ever since in much related work (Harley and Noyer 1999, Harley and Ritter 2002, Folli and Harley 2004). In DM, the lexicon is assumed to be distributed over several lists. Morphosyntactic processes derive hierarchical structures from roots and functional elements. Furthermore, morphological operations precede morphophonological realizations, by which Vocabulary Items (VI) are inserted in terminal nodes previously created by syntax. Crucially, the insertion of different VIs for the same root is dependent on properties of the subsequent syntactic context. Moreover, a realizational ("late insertion") view for both roots and functional elements is assumed (Harley 2014). Hence, allomorphy can only be triggered by elements that are linearly adjacent.

5RPDQFH WKHPH HOHPHQWV DORQJ ZLWK 3 IHDWXU a well-formedness condition on syntactic functional heads (Oltra-Massuet 1999, 2000, Pomino 2008, Pomino and Remberger 2019). More specifically, they are adjoined to little v via a node-insertion process (e.g., "node sprouting", Choi and Harley 2019). The tree structure in (1) shows the product of the syntactic derivation at spell-out while the structure in

LOOXVWUDWHV WKH DGGLWLRQDO QRGHV 7K 7KHP

8 Overall, theme vowels could be verbalizers and thus exponents of little v (Oltra-Massuet 2013, Fábregas 2017) but in Italian this seems very unlikely since in verbs such as anal-izz-a-re 'analyze' both an exponent of v ± WKH FDXVDWLYH VXIL[izz- – as well as an exponent of the theme vowel – are present simultaneously.

(1)

(2)

Following Oltra-Massuet's (1999) observations regarding Catalan conjugation classes, it is assumed that the different Italian conjugations are also hierarchically interrelated according to their degree of markedness. As discussed in §2, the first conjugation class contains mainly regular verbs and is, moreover, the most productive group, essentially thus the least marked class, followed by third conjugation verbs that present the augment who may be less regular and less productive but still more so than third conjugation verbs without augment and verbs of the second conjugation class, which are therefore both more marked.

To better understand the markedness hierarchy underlying the different conjugation classes, it thus seems plausible to think of theme elements as encoding conjugational information in terms of abstract features (Oltra-Massuet 1999, 2020). Theme elements can then be organized in a markedness hierarchy wherein those elements with a higher number of positive features are more marked than those with less or none positive features (cf. the concept of "feature geometry" in Harley and Ritter 2002). The theme markedness hierarchy is illustrated below.

(3) Theme Vowel Markedness in Standard Italian (based on Napoli and Vogel 1990)

Crucially, roots have to be minimally specified for these theme features (cf. (4)), otherwise there would be no way of ensuring that a given root matches with the theme features.

(4) (a) Roots with no specification for conjugation
¥ N D Q W
(b) Roots with specification for conjugation
¥ W H P > / @¥ S U H Q G > ¥ @L Q > @ ¥ G R U P > @

4. Issues with Previous Analyses

In Oltra-Massuet's (1999) analysis, theme allomorphy involving the Catalan variant of -isc-, i.e., -eix- (e.g., prefereixo 'I prefer'), is accounted for in terms of fusion. Fusion is a morphological device used to modify syntactic structure. It combines two sister Q R G H V L Q W R D V L Q J O H ; Û Z L W K W K H I H D W X U H V R I E A fused node thus reduces the number of terminals, providing for a single locus of vocabulary insertion (Bobaljik 2017).

In Italian, the Present Tense encodes a semantically unmarked Tense feature – just like in Catalan – and is therefore morphophonologically never realized. The structural difference between Imperfect Tense (no fusion) and Present Tense (with fusion of 7 ƒ 3 L V V K R Z Q L Q D Q G

(5) Imperfect Tense

(6) Present Tense

The same strategy is also applied in Embick (2016)'s formal analysis of Italian third conjugation verbs. According to Embick, -isc- is the phonological realization of the conjugational feature [III] in the specific context of T°[-past]. In contrast to Oltra-Massuet, Embick designates little v as the main locus of -isc- insertion (cf. (7)) though it is noted in passing that "it could just as well be treated as a Theme node" (Embick 2016, 291).

(7) Vocabulary Items for v (Embick 2016, 291)
(a) v[III] 8 -isc- / __ T[-past]
(b) v[III] 8 -i- default

To explain why the augment is not inserted in the 1 ‰ … and in the 2 ‰ …, Embick draws on another morphological strategy that is capable of changing the syntactic structure, i.e., impoverishment. The rule proposed in (8) ensures that the feature [-past] is deleted in WKH FRQWH[W RI 3> ‰ …/2 ‰ …] thus allowing only for the insertion of the default item (7b).[9]

(8) Impoverishment Rule (Embick 2016, 292)
> SDVW@ : Î BBB > SDUW SO@

In an effort to reduce the number of additional processes to account for the presence of the augment, Calabrese (2019) proposes an insertion rule that ensures the insertion of -isc- directly into roots marked with a diacritic -i in the Present Tense. This implies

[9] 2OWUD 0DVVXHW LPDSRVHV SULRU RSHUDWLRQV RQ WKH DFFRX WR DFFRXQW IRU WKH SUHVHQFH RI WKH DXJPHQW LQ WKH 3UHVHQW 6XEMXQFWLYH 3HUVRQ DQG 1XPEHU HQYLURQPHQW 2OWUD 0DVVXHW DVVXPHV WKDW WKH DXJPHQW LV LQVHUWHG DV SDUW RI WKH YRFDEXODU\ LWHPV LQ WKH ‰ … DQG LQ WKH ‰ …. However, the author argues that it is necessary to delete the whole T° node; otherwise one could not explain the absence of the augment in the Present Subjunctive 1 ‰ … and 2 ‰ …. Deleting the whole T° node in the context of 1 ‰ … and 2 ‰ … in general would lead to some sort of neutralization of the forms in this particular Person/Number environment.

that according to Calabrese, -isc- does not function as a theme element but is part of the root in certain contexts, i.e., in this case we are dealing with a case of root suppletion instead of theme allomorphy.

(9) Insertion rule (Calabrese 2019)
Insert /-isk-/ / Root_{-i}^{isk} ^ Pres

To account for the differences regarding the alternation in the 2 ‰ …, Calabrese suggests that an impoverishment rule aimed at deleting the diacritic -isk in the given environment suffices to derive the correct surface forms. Although Calabrese's approach may be appealing since fusion is avoided, the amount of additional insertion and deletion rules needed does not exactly provide for a more parsimonious formal analysis.

In the following, I propose an analysis that does indeed reduce as many post-syntactic operations as possible, in particular it will be shown that seemingly arbitrary device such as fusion is not needed.

There are many good reasons to be critical of operations such as fusion. First, it is well accepted that a simple, parsimonious model of syntactic derivation generally is to be preferred over a model that employs too powerful ad hoc devices to generate certain end structure (Trommer 1999, Haugen and Siddiqi 2016). Second, it is still unclear what factors are actually involved in triggering fusion. Finally, another big issue that has not been solved yet with regard to fusion concerns its look-ahead problem (Chung 2007, Caha 2009). Since fusion is capable of rearranging the syntactic structure it naturally has to be strictly ordered before vocabulary insertion. However, fusion only occurs when there is a portmanteau item available for insertion. No matter how we try to account for fusion it thus seems that we will be going around in circles.

But there is a solution to this problem. Fusion may be entirely unnecessary when we assume (i) a vocabulary-insertion-only model of grammar and (ii) that vocabulary insertion does not only target single terminal nodes but that two neighboring nodes may be realized by a vocabulary item at a non-terminal node (Williams 2003, Svenonius 2012, Merchant 2015, also cf. Restricted Realization Theory, as proposed by Haugen and Siddiqi 2016).

5. Allomorphy Conditioned by Post-Linearization Spanning

The core idea of the analysis is that vocabulary insertion only targets spans (Williams 2003, Melnar 2004). A span is either a head (a minimal X°) or a complement sequence of heads in a single extended projection (Svenonius 2012, cf. also Grimshaw 2000). Contrary to fusion where – due to the fact that only one VI is available for insertion (cf. §4 regarding the look-ahead problem) – two neighboring nodes combine into one, spanning enables vocabulary insertion to operate over the hierarchical structure, thus allowing to insert phonological material not just in one terminal node but also in spans

of terminal nodes that are in a complement relation with each other (cf. Williams 2003, Svenonius 2012, Merchant 2015).

In the case of fusion of features, allomorphy.[10] According to the Span Adjacency Hypothesis, however, allomorphy can only be conditioned by an adjacent span; a nonadjacent head (e.g., Z) may also cause allomorphy but only if all intervening heads (e.g., Y) are also involved in the process (see (10), cf. Merchant 2015).

(10) Span-conditioned allomorphy (based on Pomino and Remberger 2020)

Hence, a syntactic input structure such as [$_{ZP}$ Z [$_{YP}$ Y [$_{XP}$ X] Y] Z] (cf. (10)) would admit vocabulary insertion sites for <X>, <X, Y>, <Y>, <Y, Z>, <X>, <X, Y>, <X, Y, Z>, etc. (vs. the assumption that this limits the domain of contiguity to <X, Y>, or whether it could also be extended to <X, Y, Z>).

Crucially, Svenonius' notion of spanning originally refers to syntactic structure, i.e., vocabulary insertion is assumed to target syntactic structure (cf. assumptions in approaches such as Nanosyntax). But Merchant (2015) and Haugen and Siddiqi (2016) compellingly argue that a more preferable version of spanning may be one wherein (i) vocabulary insertion targets formal features rather than syntactic structure and (ii) spanning occurs post-linearization. The principle put forward by Haugen and Siddiqi (2016,369) and reproduced in (11) thus allows for the realization of strings of contiguous morphemes by a single vocabulary item (cf. also Merchant's (2015) similar Span Adjacency Hypothesis mentioned above).

---

10 Vocabulary insertion proceeds from inside out, i.e., starting with the root. The process is guided by the principle of inward vs. outward sensitivity which means Vocabulary insertion at node Y can see the phonological and morphological features already exponed at node X but it can only see the morphosyntactic features of a not yet realized node Z (Bobaljik 2000, Embick 2010).

(11) Post-linearization contiguous insertion principle
Insertion may realize multiple adjacent X°'s (features) provided that the features realized by the inserted Vocabulary item are as large a subset of the string of adjacent X°'s (features than that which could otherwise be expressed by separate overt Vocabulary items at the contained X°'s (features). (Haugen and Siddiqi 2016, 369–370)

As anillustration of this principle we can now turn to the theme allomorphy we find in augmented Italian -isc- verbs.

In the Present Indicative and Subjunctive the underlying phonological vocabulary item /isk/ is realized as /iskj/ in the 1Pl and 2Pl forms and as /isk/ in all other persons. Elsewhere, the theme vowel is realized as /ja/ (1‰ …, 2‰ … Subjunctive) and /i/ (2‰ … Indicative) instead (cf. Table 1). With regard to the principle of post-linearization spanning, a terminal node such as 3Pl Fut can realize an adjacent span comprising <Th[III], T°[-past], Th[Î ?>.

(12) Terminal Span 3Pl Fut conditions allomorphy as follows: /isk/ ⇔ <Th[III], T°[-past], Th[Î ?>

One major concern that may arise with regard to (12) is the nature of the span that comprises features that are structurally not adjacent to each other, i.e., Th[III], T°[past] and Th[Î ?. However, the idea of post-linearization spanning is precisely suited to account for this data. It is crucial to keep in mind that the portmanteau formed here concerns linearly adjacent nodes, without reference to their structural constituency (cf. also Ostrove 2018). This is a feasible approach considering that the Present Tense is a phonologically unmarked tense in Italian. Since the theme allomorphy with -isc- only occurs in the Present Tense, we may thus assume that a span comprising conjugational features and tense features does not occur in other tenses where T exhibits an autonomous phonological exponent (cf. (5), e.g., 3‰ … Imperfect finì-va-no 'they finished' with /i/ being the theme vowel of the base and /a/ representing the theme vowel of the Imperfect Tense). In these cases, the VIs at our disposal have different "span sizes" as well as different feature specifications.

Based on these observations, we can identify the following spans for vocabulary insertion, depending on the syntactic context:

(13) VIs for <Th, T°, Th> (Present Tense)

D 7 K > , , , @ 7 ƒ > S D V W V E M @Œ € 7 ‰ > Î] > 8 L V N L L I 3

E 7 K > , , , @ 7 ƒ > S D V W V E M @‰ … 7 K > Î] > 8 M D L I 3

(c) <Th[III], T°[-past], Th[Î @ ! < M D ‰ L … > 3

(d) <Th[III], T°[-past], Th[Î]> 8 L V N L Œ € 3 ‰ …>

(e) <Th[III], T°[-past], Th[Î]> 8 L U U L Œ € > 3

(f) <Th[III], T°[-past], Th[Î]> 8 /i/

The items in (13) are ordered in terms of decreasing complexity but also according to language specific relations among different features involved in the competition (Halle and Marantz 1993, 150). For instance, the featural composition of the spans in (13a) and (13b) is more complex (and the context of insertion is also more specific) than the one in (13c–e); the least complex environment of insertion is (13f) which ultimately presents the default theme exponent that actually also occurs in other tenses.

7 K H W U L Y L D O V S D Q V U H S U H V H Q W L Q J W K H 3 I H D W X U H

(14) 9 , V I R U 3 !

(a) <1 ‰ … ! < P R

(b) <2 ‰ … ! < W H

(c) <3 ‰ … ! < Q R

(d) <1 Œ € ! < R

(e) <2 Œ € ! < L

(f) There is no exponent for Œ €.

Now, turning back to our derivation at hand, Spell-Out occurs first. The syntactic structure that is being generated at this point can be found in (15), reproduced from (1). Our morphological starting point is however (16), reproduced from (2), wherein additional W K H P H Q R G H V D Q G D 3 Q R G H K D Y H E H H Q D G G H G S R V morphological well-formedness.

(15)

(16)

We know that the nodes in (16) will be linearized as illustrated in (17a).

(17) D / L Q H D U L] D W L R Q ¥ U R R W -past >@ @ 3 > , , , @ > S D V W @

(b) Insertion: /fin/ /iski/ /o/

(c) Readjustment: /finisko/ (vowel deletion)

As illustrated in (17), there are still some phonological readjustment rules that we need to apply after the vocabulary insertion process (e.g., vowel deletion in 1 Œ €, cf. (17b–c)) but, contrary to morphological readjustment rules and additional operations such as fusion, these phonological rules comply with general diachronic observations regarding the Italian vowel system. We may propose alternative phonological exponents that already represent the final phonological surface form (e.g., /isko/ for the 3 ‰ … L U U H I R U but we would then miss the unifying link between these verb forms, i.e., the main X Q G H U O \ L Q J I R U P V L V N L D Q G L U U L I U R P Z K L F K Z diachronically plausible way.

6. Conclusions

The analysis proposed in this paper provides empirical support for some aspects of Haugen and Siddiqi's (2016) Restricted Realization Theory. In particular, it has been shown that the traditional process of vocabulary insertion via terminal nodes may benefit from non-terminal insertion via post-linearization spanning. Theme allomorphy, as it was discussed in this paper, may be explained via spans of features wherein featural adjacency is more relevant than structural adjacency.

With regard to previous analyses involving the Italian augment -isc-, the analysis presented here manages to hold without having to resort to additional morphological operations such as fusion and impoverishment. Overall, this new analysis of the Italian theme formative -isc- may hopefully give further impetus to advancing certain theoretical conceptions in DM.

Works Cited

Anderson, Stephen R. 1992. A-Morphous Morphology. Cambridge: Cambridge University Press.

Aronoff, Mark. 1994. Morphology by Itself. Stems and Inflectional Classes. Cambridge, MA: MIT Press.

Arregi, Karlos. 2000. “How the Spanish Verb Works.” Presented at the 30th Linguistic Symposium on Romance Languages, University of Florida, Gainesville. Accessed October 14, 2021. http://home.uchicago.edu/~karlos/Arregi-2000-how.pdf.

Bobaljik, Jonathan David. 2000. “The Ins and Outs of Contextual Allomorphy.” In Proceedings of the Maryland Mayfest on Morphology 1999, edited by Kleanthes Grohmann and Caro Struijke, 35–71. University of Maryland Working Papers in Linguistics 10. College Park: Department of Linguistics, University of Maryland.

Bobaljik, Jonathan David. 2017. “Distributed Morphology.” Oxford Research Encyclopedia of Linguistics (Online Encyclopedia), edited by Mark Aronoff. Oxford: Oxford University Press. Accessed October 14, 2021. https://doi.org/10.1093/acrefore/9780199384655.013.131.

Caha, Pavel. 2009. “The Nanosyntax of Case.” PhD diss., University of Tromsø.

Calabrese, Andrea. 2019. “Morphophonological Investigations: A Theory of PF. From Syntax to Phonology in Italian and Sanskrit Verb Forms.” Unpublished Ms., University of Connecticut.

Choi, Jaehoon, and Heidi Harley. 2019. “Locality Domains and Morphological Rules: Phases, Heads, Node-Sprouting and Suppletion in Korean Honorification.” Natural Language & Linguistic Theory 37 (4): 1319–65.

Chung, Inkie. 2007. “Suppletive Negation in Korean and Distributed Morphology.” Lingua 117 (1): 95–148.

Da Tos, Martina. 2013. “The Italian FINIRE-Type Verbs: A Case of Morphomic Attraction.” In The Boundaries of Pure Morphology, edited by Silvio Cruschina, Martin Maiden, and John Charles Smith, 45–67. Oxford: Oxford University Press.

Dressler, Wolfgang U., and Anna Thornton. 1991. “Doppie Basi e Binarismo Nella Morfologia Italiana.” Rivista Di Linguistica 3: 3–22.

Embick, David. 2010. Localism versus Globalism in Morphology and Phonology. Cambridge, MA: The MIT Press.

Embick, David. 2016. “On the Distribution of Stem Alternants: Separation and Its Limits.” In The Morphome Debate: Diagnosing and Analyzing Morphomic Patterns, edited by Ricardo Bermúdez-Otero and Ana Luís, 276–305. Oxford: Oxford University Press.

Fábregas, Antonio. 2017. “Theme Vowels Are Verbs.” Accessed October 14, 2021. https://lingbuzz.com/j/rgg/2017/2017.05/f%C3%A1bregas_theme-vowels-are-verbs_RGG-2017-05.pdf.

Folli, Raffaella, and Heidi Harley. 2005. “Flavors of v.” In Aspectual Inquiries, edited by Paula Kempchinsky and Roumyana Slabakova, 95–120. Studies in Natural Language and Linguistic Theory. Dordrecht: Springer.

Grimshaw, Jane. 2000. “Locality and Extended Projection.” In Current Issues in Linguistic Theory, edited by Peter Coopmans, Martin B.H. Everaert, and Jane Grimshaw, 115–134. Amsterdam: John Benjamins Publishing Company.

Halle, Morris, and Alec Marantz. 1993. “Distributed Morphology and the Pieces of Inflection.” In The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger, edited by Ken Hale and Samuel Jay Keyser, 111–76. Cambridge, MA: MIT Press.

Halle, Morris, and Alec Marantz. 1994. “Some Key Features of Distributed Morphology.” MIT Working Papers in Linguistics 21: 275–88.

Harley, Heidi. 2014. “On the Identity of Roots.” Theoretical Linguistics 40 (3–4): 225–276.

Harley, Heidi, and Rolf Noyer. 1999. “State-of-the-Article: Distributed Morphology.” Glot International 4 (4): 3–9.

Harley, Heidi, and Elizabeth Ritter. 2002. “Person and Number in Pronouns: A Feature-Geometric Analysis.” Language 78: 482–526.

Haugen, Jason D., and Daniel Siddiqi. 2016. “Towards a Restricted Realization Theory: Multimorphemic Monolistemicity, Portmanteaux, and Post-Linearization Spanning.” In Morphological Metatheory, edited by Daniel Siddiqi and Heidi Harley, 343–86. Amsterdam: John Benjamins.

Julien, Marit. 2015. “Theme Vowels in North Sámi: Spanning and Maximal Expression.” Lingua 164: 1–24.

Maiden, Martin. 2004. “Verb Augments and Meaninglessness in Early Romance Morphology.” Studi Di Grammatica Italiana 22: 1–61.

Maiden, Martin. 2005. “La Ridistribuzione Paradigmatica Degli “Aumenti” Verbali Nelle Lingue Romanze.” In Latin et Langues Romanes. Études de Linguistique Offertes à József Herman à l’occasion de Son 80 Ème Anniversaire, edited by Sandor Kiss, Luca Mondin, and Giampaolo Salvi, 431–40. Tübingen: Niemeyer.

Maiden, Martin. 2009. “From Pure Phonology to Pure Morphology the Reshaping of the Romance Verb.” Recherches Linguistiques de Vincennes 38: 45–82.

Maiden, Martin. 2018. The Romance Verb. Morphomic Structure and Diachrony. Oxford: Oxford University Press.

Melnar, Lynette. 2004. Caddo Verb Morphology. Lincoln, NB: University of Nebraska Press.

Merchant, Jason. 2015. “How Much Context Is Enough? Two Cases of Span-Conditioned Stem Allomorphy.” Linguistic Inquiry 46 (2): 273–303.

Meul, Claire. 2010. “The Intra-Paradigmatic Distribution of the Infix -l/ESC- from Latin to Modern Romance: Morphomic Patterning and Beyond.” Morphology 20 (1): 1–40.

Meul, Claire. 2013. 7 K H 5 R P D Q F H 5 H O H [H V R I W K H - / 5 W V W Q U X F W L [Turing and Remodeling Processes. Hamburg: Buske.

Napoli, Donna Jo, and Irene Vogel. 1990. “The Conjugations of Italian.” Italica 67 (4): 479–502.

Oltra-Massuet, Isabel. 2020. “Conjugation Class.” In Oxford Research Encyclopedia of Linguistics (Online Encyclopedia), edited by Mark Aronoff. Oxford: Oxford University Press. Accessed October 14, 2021. https://doi.org/10.1093/acrefore/9780199384655.013.545

Oltra-Massuet, Maria Isabel. 1999. “On the Constituent Structure of Catalan Verbs.” In Papers in Morphology and Syntax, Cycle One. MIT Working Papers in Linguistics 33, edited by Karlos Arregi, Benjamin Bruening, Cornelia Krause, and Vivian Lin, 279–322.

Ostrove, Jason. 2018. “Stretching, Spanning, and Linear Adjacency in Vocabulary Insertion.” Natural Language and Linguistic Theory 36 (4): 1263–89.

Pomino, Natascha. 2008. Spanische Verbalflexion. Eine minimalistische Analyse im Rahmen der Distributed Morphology. Tübingen: Niemeyer.

Pomino, Natascha, and Eva-Maria Remberger. 2019. “Verbal Suppletion in Romance Synchrony and Diachrony: The Perspective of Distributed Morphology.” Transactions of the Philological Society. Special Issue: The Diachrony of Suppletion 117 (3): 471–97.

Scalise, Sergio. 1994. Morfologia. Bologna: Il Mulino.

Schwarze, Christoph. 1999. “Inflectional Classes in Lexical Function Morphology: Latin -sk- and Its Evolution.” In Proceedings of the LFG 99 Conference, University of Manchester, edited by Tracy Holloway King and Miriam Butt, 1–13. Stanford: CSLI Publications.

Svenonius, Peter. 2012. “Spanning.” Ms., Center of Advanced Study in Theoretical Linguistics, University of Tromsø. Accessed October 14, 2021. https://ling.auf.net/lingBuzz/001501.

Trommer, Jochen. 1999. “Morphology Consuming Syntax’ Resources: Generation and Parsing in a Minimalist Version of Distributed Morphology.” In Proceedings of the ESSLI Workshop on Resource Logics and Minimalist Grammars, edited by Christian Retoré and Edward Stabler. Nijmegen: Radboud Universiteit Nijmegen. Accessed October 14, 2021. http://www.ai.mit.edu/projects/dm/trommer-consuming.pdf.

Vincent, Nigel. 1988. “Italian.” In The Romance Languages, edited by Martin Harris and Nigel Vincent, 279–313. London: Croom Helm.

Williams, Edwin. 2003. Representation Theory. Current Studies in Linguistics Series 39. Cambridge, MA: MIT Press.

Head Movement in Germanic Doubly-Filled Comp Constructions

Dalina Kallulli [a] and Sabine Laszakovits [b,c]

[a]University of Vienna, Vienna, Austria; [b]Austrian Academy of Sciences, Vienna, Austria; [c]University of Connecticut, Storrs, CT, USA

[a]dalina.kallulli@univie.ac.at; [b]sabine.laszakovits@oeaw.ac.at

Abstract: Taking our cue from Bayer and Brandner (2008), in this paper we investigate the feature composition of so-called (morphologically) “simplex” wh-elements in German, e.g. wer “who” or was “what”, versus “complex” wh-elements with internal structure, e.g. “how many” or “for which student”. We show that simplex wh-elements do not contain any features other than [+wh]; in particular, they do not contain the features D (including its sub-features Person, Number, and Gender), Case, nor Animacy. This approach makes correct predictions on at least two accounts: (i) headless relatives are headed by simplex wh-words, and not by relative pronouns, due to the lack of features of the antecedent; and (ii) combined with Roberts’ (2010) theory of head movement, we derive the data of Bavarian and Alemannic doubly-filled complementizer constructions reported in Bayer and Brandner (2008) for embedded questions, which despite (possibly cross-generational) speaker variation, crucially also represent our own grammar, and we also extend our analysis to doubly-filled complementizer effects in relative clauses in the same varieties.

Keywords: German; doubly-filled complementizer; wh-words; relative pronouns; r-pronouns

1. Introduction

South German varieties, such as Bavarian and Alemannic, display Doubly-Filled Comp in some but not all embedded questions, as shown in (1a) versus (1b), and in relative clauses, as shown in (2):

1 We thank Ian Roberts for his close reading and detailed comments, and the OLINCO reviewers and audience for their questions.

(1) (a) I frog mi fia wos dass =a des kauft hot.
I ask myself for what that =he this bought has
“I wonder why he bought this.”

(b) I frog mi wen (*dass) =a J V H ¾ hot.
I ask myself who that =he seen has
“I wonder who he has seen.”

(2) Da Tisch, den wos i kauft hob, gfoit da Mama.
the. ‡ ˆ †table ‹ ~ … Œ €wh zat | I bought have pleasesthe. } z• mom
“Mom likes the table that I have bought.”

Concentrating on their observation of the contrast between data like (1a) and (1b), Bayer and Brandner (2008) argue that wh-elements fall into two classes, namely: (i) class 1, which contains morphologically complex wh-elements, such as wia-vü “how much/many”, wa-rum “what-for” (i.e. “why”), mit wem “with whom”, which appear with dass “that” in C, as shown in (1a); and: (ii) class 2 wh-elements, which are morphologically simplex, such as wos “what”, wer “who. ‡ ˆ Wen “who. z | |”, wia “how”, wo “where” etc. In the variety of Bayer and Brandner as in ours, the latter (i.e., class 2 wh-elements) are incompatible with dass, as shown in (1b).

Bayer and Brandner (2008) analyze this contrast to mean that simplex (i.e., class 2) wh-elements occupy C0, thus blocking the insertion of dass, which is in their account further corroborated by evidence from clitic placement (namely, the so-called Wackernagel position) and clitic phonology, namely sandhi and r-epenthesis: specifically, simplex but not complex wh-elements show sandhi and r-epenthesis effects at vowel hiatus with clitics, as illustrated in (3) through[2] (5)

The examples in (3) show that the simplex wh-word wer “who” displays sandhi before vowels (3a), but not before consonants (3b), and the complex wh-word welch-einer “which one” doesn’t show any sandhi, even before vowels (3c).

(3) (a) I frog mi wer [vea-r] =i bin.
I ask myself who =I am
“I wonder who I am.”

(b) I frog mi wer [vea] =ma san.
I ask myself who we are
“I wonder who we are.”

2 While stating the observation, Bayer and Brandner (2008) do not illustrate sandhi and H S H Q W K H V L V H u H F W V W K H H [D P S O H V L Q W K U R X J K D

(c) *I frog mi welch-einer [vø:ç-a:na]-a =a is.
I ask myself which-one =he is
Intended: “I wonder which one he is.”

The examples in (4) and (5) illustrate r-epenthesis between vowels. In (4a) and (4b), we see r-epenthesis between the simplex wh-words wo “where” and wia “how” and the vowel-initial clitic. In (5a) and (5b) we see that r-epenthesis is not possible with the complex wh-words von wo “from where” and wia-vü “how much”.

(4) (a) I frog mi wo-[r] =a woa.
I ask myself where =he was
“I wonder where he was.”

(b) I frog mi wia-[r] =a des gmocht hot.
I ask myself how =he this done has
“I wonder how he did this.”

(5) (a) *?I frog mi von wo-[r] =a herkommt.
I ask myself from where =he comes
Intended: “I wonder where he comes from.”

(b) *I frog mi wia-vü-[r] =a kauft hot.
I ask myself how-much =he bought has
Intended: “I wonder how much he bought.”

Bayer and Brandner’s (2008) findings are summarized in Table 1.

	Simplex wh-elements	Complex wh-elements
Dass-insertion in embedded questions	No	Yes
Sandhi with clitic pronouns	Yes	No
R-epenthesis with clitic pronouns	Yes	No

Table 1. Bayer and Brandner (2008)

The main theoretical problem with Bayer and Brandner’s (2008) account, as well as Bayer’s (2014), is that they propose a dubious re-Merge operation: a lexical item can contain an (effectively) Greedy selection feature, and after this item has entered the derivation, the Greedy feature will wait for a predefined context to arise and will then trigger its

re-Merge by selecting the derivation's root node. Concretely for the cases at hand, this works in the following way. Wh-words contain a greedy ("latent") feature [?C], which is on the lookout for a configuration where the wh-word can become the sister of TP. When this possibility arises, this feature triggers to re-Merge the wh-word as C-head, projecting a CP. This possibility arises for simplex wh-words, but not for complex ones, as they are embedded inside their own phrase. Thus, complex wh-words do not become C-heads.

Empirically, Bayer and Brandner's (2008) account says nothing about the fact that the same phenomenon (i.e. Doubly-Filled Comp) is in the same varieties found not only in embedded questions with complex wh-elements, but also in relative clauses, as was shown in (2). Furthermore, unlike embedded questions, relative clauses always allow C^0 insertion, with C^0 taking the form of wos or wo, depending on the variety (Bayer 1984). Thus, a core observation we make, is that the simplex wh-word wen in (1b) and the relative pronoun den in (2) behave differently in this respect.

In what follows, we provide a uniform and unified account of the phenomenon of Doubly-Filled Comp in both embedded questions and relative clauses in these varieties of German, which also evades the theoretical problems (namely, Greed and re-Merge by selecting the derivation's root node) of this previous approach.

2. Proposal

The analysis that we propose for the data described in the previous section contains three core ingredients. Firstly, we argue that simplex wh-elements are underspecified in terms of their feature constitution. Secondly, we assume Roberts' (2010) theory of head movement. Thirdly, we assume default spell-out rules arguing in particular that the default spell-out of empty D^0 in the varieties of German that exhibit the Doubly-Filled Comp in embedded questions is dass. We now turn to each of these three ingredients.

2.1 The Feature Constitution of Wh-Words

We argue that wh-elements in Bavarian and other Southern German varieties have no features other than [+wh]. In particular, they do not contain the features Case, Animacy, nor D (including its sub-features Person, Number, and Gender).

As far as Case is concerned, we follow Marantz (1991) and many others who take Case to be assigned post-syntactically (i.e., Case is not a feature that is present early in the derivation)[3].

---

[3] Spezifisch sogar syntaktisch ist die Idee, dass Kasus in drei Schritten zugewiesen wird:

(i) Assign lexical cases (as governed by verbs or adpositions);
(ii) Assign dependent cases (based on structural relations between arguments);
(iii) Assign unmarked and default cases.

Furthermore, we argue that wh-elements do not encode animacy, and more specifically that animacy is actually not grammaticalized in German. Instead, we follow Wurmbrand (2017) in that each noun has a formal gender feature out of the set {†, •, ‡} (for masculine, feminine, and neuter, respectively), and may have a semantic gender feature out of the set {†, •}. Inanimacy is then simply non-animacy: Inanimate nouns are characterized by the absence of a semantic gender feature. In this way, animacy falls out of the gender system without the need to postulate an independent feature [±animate]. This approach derives the following two constructions that could potentially be understood as counter-arguments to the absence of animacy in the syntax, namely (i) gender mismatches and (ii) so-called "R-pronouns" (van Riemsdijk 1978), to which we turn next.

There seems to exist an animacy/inanimacy split in gender mismatch constructions. For instance, the noun Mädchen "girl" has the formal gender ‡ and the semantic gender •. A possessive pronoun agreeing with Mädchen can however appear with either feature, as shown in (6).

(6) Das Mädchen sitzt an {⁹seinem / ⁹ihrem} Platz. ⁹formal/⁹semantic
the.‡ girl(‡) sits at its.‡ Œher.• Œspot
"The girl is sitting at her seat."

In contrast, as shown in (7), an inanimate noun like Gabel "fork" with formal gender • can only take possessive pronouns that agree with it with ‡ as might be expected given that a fork is a thing.

(7) Die Gabel liegt an {⁹ihrem / *seinem} Platz. ⁹formal/*semantic
the.• fork(•) lies at her.• Œits.‡ Œspot
"The fork is lying in its spot."

We argue that the absence of neuter agreement for Gabel "fork" in (7) is not due to this word's inanimacy, but to Gabel "fork" not having a semantic gender feature.[4]

Turning to R-pronouns (such as e.g. darauf "there on"/"thereafter", darüber "above"/"across"/"furthermore", etc.), these expressions in German can famously refer to inanimate nouns, but not to animate nouns.[5] We argue that the class of expressions that R-pronouns can refer to is correctly characterized not by inanimacy, but rather by non-animacy, which derives why R-pronouns may not only refer to inanimate nouns,

4 This argument is analogous to the discussion of formal versus semantic number features in Hiaki (Harley 2019).

5 As to not unnecessarily gendered [sic] language, we use the gender-neutral pronoun hir (Hoelig Fleischer (2002), among many others.

but also to infinitives (8a), CPs (8b), and events (8c), none of which are characterizable in terms of (in)animacy features.

(8) (a) Ich warte darauf, ihn morgen zu sehen.
I wait there-on him tomorrow to see. , ‡ •
“I am waiting to see him tomorrow.”

(b) Ich warte darauf, dass ich ihn morgen sehe.
I wait there-on that I him tomorrow see. Œ €
“I am waiting that I see him tomorrow.”

(c) Er hat ihr gratuliert. Darüber freut sie sich.
he has her congratulated there-over is.happy she ‹ ~ • …
“He congratulated her. She is happy about that.”

Finally, we argue that wh-elements do not contain the feature D when they are not D-linked. In this, they differ from D-linked wh-elements, which we argue to contain D, building on Pesetsky (1987) and Heck and Müller (2000). Along with many others, notably Richards (2004, 2008) and Bárány (2017), we understand the features Person, Gender, and Number to be sub-features of D, and therefore to be absent in non-D-linked wh-elements.

As we discuss in more detail in section 2.2, the third person singular agreement of simplex wh-words such as wer “who” and was “what” with verbs and pronouns is not agreement triggered by a Person or Number probe, but just default agreement. Evidence for this abounds: infinitives (9a), finite CPs (9b), and impersonal passives (9c) also trigger third person singular neuter agreement.

(9) (a) Genug zu schlafen ist wichtig.
enough to sleep. , ‡ • is. Œ € important
“It is important to get enough sleep.”

(b) Dass du gekommen bist, hat auch seinen Vorteil.
that you come.‰ ‰ ‰ are has. Œ € also its. Œ € advantage
“There is also an advantage to your having come.”

(c) Im Saal wird getanzt.
in.the ballroom becomes. Œ € danced‰ ‰ ‰
“People are dancing in the ballroom.”

That wh-words are featureless is evidenced from their appearance in headless relative clauses. While in headed relative clauses as in (10) the relative pronoun carries the features Number and Gender as copied from the head noun, in headless relatives the appearance of the relative pronoun d- is impossible even with the default features for animates† Œ € Instead, a wh-word appears, as in (11).

(10) Ich mag den Arzt, { ⁹den / *wen } sie mag.
I like the. † doctor(†) ‹ ~ … †whŒ € she likes
"I like the doctor she likes."

(11) Ich mag, {*den / ⁹wen } sie mag.
I like ‹ ~ … † whŒ € she likes
"I like who(ever) she likes."

We now turn to the second core ingredient of our analysis, namely syntactic head movement, as developed in Roberts (2010).

2.2 Roberts' (2010) Theory of Head Movement

Roberts (2010) develops a syntactic theory of head movement, which we adopt wholesale. Specifically, Roberts (2010) argues that for a constituent Ù to attach to Ú via head-movement, two conditions need to be met, namely: (i) Ù must be simultaneously minimal and a maximal projection; and (ii) Ù's features must be a subset of Ú's features.

With this background, let us now see how we derive (1b), repeated here for ease of reference:

(1) (b) I frog mi wen (*dass) =a J V H ¾ hot.
I ask myself who that =he seen has
"I wonder who he has seen."

In (1b), the embedded C is C[wh,fin]. As we have argued above, wen "who" has exactly the feature set [wh], and it does not have internal structure. This means that both of Roberts' conditions for head movement are fulfilled: [wh] is a subset of [wh,fin]. The derivation would thus include the following steps: (i) C needs to check its [wh] feature; (ii) C attracts a constituent with [wh], namely wen; (iii) wen moves into C via head-movement thus blocking the presence of the overt complementizer dass "that" and turning C into a complex head, as shown in (12):

(12)

In contrast, with complex wh-elements there can be no head movement to C and only phrasal movement is possible. Thus, in (1a), repeated below for ease of reference, the embedded C is C[wh,fin] and fia wos "for what" is a complex wh-element with internal structure since the wh-element in it is the complement of a prepositional head, which means that Roberts' (2010) first condition for head movement is not fulfilled since the mover is not a minimal category; see the structures in (13) and (14).

(1) (a) I frog mi fia wos dass =a des kauft hot.
I ask myself for what that =he this bought has
"I wonder why he bought this."

(13)

(14)

6 Nota bene that movement of wos alone, stranding ¿ D, is not possible in general since German does not have English-style preposition stranding. The only preposition-stranding available in German occurs with R-pronouns, and only in some German dialects, crucially not in the ones discussed here.

We also correctly derive that simplex wh-words undergoing successive-cyclic movement do not land in C of intermediate landing sites, since these Cs' features are not a superset of [+wh].

Let us now turn to Doubly-Filled Comp in relative clauses as in (2), repeated here for ease of reference.

(2) Da Tisch den wos i kauft hob gfoit da Mama.
the.‡ˆ†table ‹~… Œ€whatˆ| I bought have pleases the.}z• mom
"Mom likes the table that I have bought."

Relative pronouns such as den in (2) agree with their antecedent in gender and number. The C in a relative clause, namely wos in (2), has the feature [rel] but certainly not the features gender and number since this is invariable no matter what the gender and number of the head noun is. Head movement of den to the embedded C is not possible because its features are not a subset of the features of the attachment site, which means that Roberts' (2010) second condition for head movement is violated. More specifically, the derivation in this case includes the following steps: (i) C[rel] attracts the relative pronoun den, and (ii) den undergoes phrasal movement to Spec,CP, as given in (15).

(15)

2.3 Default Spell-Out of Empty C

We propose that Bavarian (and similar varieties such as Alemannic) differ from Standard German in a micro-parameter (Roberts 2019) setting, namely: (i) Bavarian/Alemannic empty C[wh] is spelled out as dass; (ii) Bavarian/Alemannic empty C[rel] is spelled out as wos or wo; (iii) Standard German empty C[wh] and empty C[rel] remain unpronounced.

The motivation for this involves the following reasoning. While Standard German does not have clitic pronouns (Cardinaletti and Starke 1999), in Bavarian and Alemannic many pronouns are clitics that attach themselves, or move, to C, which is traditionally known as the Wackernagel position. When clitics attach to a host, the host must be phonologically overt. Thus, C must be overt in contexts with clitics for the derivation to succeed. We propose that C is overt not only in these cases, but has been grammaticalized

into being overt. Simplex wh-words fill C themselves, and they can host clitics, as the presence of sandhi (3a) and r-epenthesis (4) evidence (as reported in Bayer and Brandner 2008). Complex wh-words on the other hand move to Spec,CP, and they cannot host clitics, as witnessed by the absence of sandhi (3c) and r-epenthesis (5) (see again Bayer and Brandner 2008). In these cases, C is filled by dass or was or wo, depending on C's features and the variety.

3. Conclusion

Our aim was to provide an account of doubly-filled complementizer effects observed in embedded questions and relative clauses in Bavarian and other Southern German varieties. We have followed Bayer and Brandner's (2008) observation about embedded questions, particularly that there are two classes of wh-words with different syntactic properties and that only the simplex ones (which are located in C), but not the complex ones (which are located in Spec,C) show doubly-filled complementizer effects, and we have provided a simpler account of these data. Our analysis relies on and provides additional empirical coverage for Roberts' (2010) theory of head movement, which is based on two conditions, namely no internal structure of the mover, and features of the mover being a subset of the features in the landing site. We have shown that complex wh-elements violate Roberts' I L U V W F R Q G L W L R Q I R U K H D G P R Y H P H Q W D Q G W K D W violate Roberts' second condition, and are thus unable to undergo head movement. We have furthermore proposed that if no head movement happens, C is parameterized to be filled by default withdass or was/wo in order to host clitics, which unlike in Standard German are as pervasive in these varieties as the doubly-filled comp. Further scrutiny notwithstanding, our analysis can also account for similar constructions found in other Germanic languages, such as in Swiss German embedded questions and relative clauses (van Riemsdijk 1989), in Yiddish relative clauses (Lowenstamm 1977), and even in Standard English relative clauses (Pesetsky and Torrego 2006).

References

Bayer, Josef. 1984. "COMP in Bavarian Syntax." The Linguistic Review í

Bayer, Josef. 2014. "Syntactic and phonological properties of wh-operators and wh-movement in Bavarian." In Bavarian Syntax. Contributions to the Theory of Syntax [Linguistik Aktuell/Linguistics Today 220], edited by Günther Grewendorf D Q G + H O P X W : H L V V í $ P V W H U G D P - R K Q % H Q

Bayer, Josef, and Ellen Brandner. 2008. "On Wh-Head-Movement and the Doubly-Filled-Comp Filter." In Proceedings of the 26th West Coast Conference on Formal Linguistics H G L W H G E \ & K D U O H V % & K D Q J D Q G + D Q Q D K Cascadilla Proceedings Project.

Bárány, András. 2017. Person, Case, and Agreement: The Morphosyntax of Inverse Agreement and Global Case Splits. Oxford: Oxford University Press.

Cardinaletti, Anna and Michal Starke. 1999. “The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns.” In Clitics in the Languages of Europe H G L W H G E \ + H Q N Y D Q 5 L H P V G L M N í 0 R X
Fleischer, Jürg. 2002. Die Syntax von Pronominaladverbien in den Dialekten des Deutschen. Eine Untersuchung zu Preposition Stranding und verwandten Phänomenen. Franz Steiner Verlag.
Harley, Heidi. 2019. “Hiaki Root Number Suppletion.” Talk given at the 12th Brussels Conference on Generative Linguistics: Suppletion, Allomorphy, and Syncretism % & * / R Q ' H F í D W . 8 / H X Y H Q % H O J L X P
Heck, Fabian and Gereon Müller. 2000. “Successive Cyclicity, Long-Distance Superiority, and Local Optimization.” In WCCFL 19 Proceedings, edited by Roger Billerey and % U R R N ' D Q L H O O H / L O O H K D X J H Q í 6 R P H U Y L O O
Helbig, Gerhard. 1974. “Bemerkungen zu den Pronominaladverbien und zur Pronominalität.” Deutsch als Fremdsprache í
Lowenstamm, Jean. 1977. “Relative Clauses in Yiddish: A Case for Movement.” Linguistic Analysis í
Marantz, Alec. 1991. “Case and Licensing.” In Proceedings of the 8th Eastern States Conference on Linguistics (ESCOL 8), edited by German Westphal, Benjamin Ao and Hee-Rahk Chae, 234–253. Ithaca, NY: CLC Publications.
Pesetsky, David. 1987. “Wh-in-situ: Movement and Unselective Binding.” In The Representation of (In)definiteness H G L W H G E \ (U L F 5 H X O D Q G D Q G $ C MIT Press.
Pesetsky, David and Esther Torrego. 2006. “Probes, Goals and Syntactic Categories.” In Proceedings of the 7th annual Tokyo Conference on Psycholinguistics, edited by Yukio Otsu. Tokyo: Hituzi Syobo Publishing Company.
Richards, Marc. 2004. “Object Shift, Scrambling, and Symmetrical Syntax.” PhD diss., University of Cambridge.
Richards, Marc. 2008. “Defective Agree, Case Alternations, and the Prominence of Person.” In Scales (Linguistische Arbeitsberichte 86), edited by Marc Richards and Andrej L. Malchukov, 137–61. Leipzig: Universität Leipzig.
Roberts, Ian. 2010. Agreement and Head Movement. Clitics, Incorporation, and Defective Goals. Cambridge, MA: MIT Press.
Roberts, Ian. 2019. Parameter Hierarchies and Universal Grammar. Oxford: Oxford University Press.
van Riemsdijk, Henk. 1978. A Case Study in Syntactic Markedness. The Binding Nature of Prepositional Phrases. Dordrecht: Foris.
van Riemsdijk, Henk. 1989. “Swiss Relatives.” In Sentential Complementation and the Lexicon. Studies in Honor of Wim De Geest H G L W H G E \ ' D Q \ - D V S H U V < Y D Q 3 X W V H \ D Q G 3 L H W H U 6 H X U H Q í ' R U G U H

Wurmbrand, Susi. 2017. “Formal and Semantic Agreement in Syntax: A Dual Feature Approach.” In Language Use and Linguistic Structure: Proceedings of the Olomouc Linguistics Colloquium 2016, edited by Joseph Emonds and Markéta Janebová, í 2 O R P R X F 3 D O D F N ê 8 Q L Y H U V L W \

The Link between Athematicity and Irregularity in French Verbal Inflection

Natascha Pomino [a] and Eva-Maria Remberger [b]

[a]University of Wuppertal, Wuppertal, Germany and [b]University of Vienna, Vienna, Austria

[a]pomino@uni-wuppertal.de and [b]eva-maria.remberger@univie.ac.at

Abstract: This paper takes Romance conjugation classes as derived from Latin as its starting point. We claim that theme vowels are not only retained in Spanish, Italian etc., but also in French: A reflex of athematic forms is found in latent root-final consonants, such as in forms of vivre ‘to live’, whereas in thematic classes these consonants are saved by the presence of a theme vowel, as in forms of arriver ‘to arrive’. Based on this assumption, we propose a feature geometry for French conjugation classes, including neutralization processes for certain categorial contexts. Using a Spanning approach within the framework of Distributed Morphology, we demonstrate the accuracy of the Suppletion Generalization (Vanden Wyngaerd 2018), which states that irregularity of form is dependent on a reduced number of affixes, not only in the case of missing theme vowel positions, but also in the case of non-categorial suppletion.

Keywords: Spanning, suppletion, allomorphy, French, theme vowels

1. Introduction

Work on systematicity in suppletion and stem allomorphy has shown that irregularity in the forms of inflected elements is related to a reduced number of affixes in these forms (cf. Bobaljik 2012, Caha et al. 2019). Following this line of reasoning, Vanden Wyngaerd (2018) proposes the Suppletion Generalization in (1) which he illustrates, among other things, with the Italian past participles given in Table 1.

(1) Suppletion Generalization: If there is irregularity in the form of either the root or the suffixes, the number of suffixes gets reduced (Vanden Wyngaerd 2018, 1).

regular					irregular				
am-	-a-	-t-	-o	‘loved’	per-	Ø	-s-	-o	‘lost’
batt-	-u-	-t-	-o	‘beaten’	cor-	Ø	-s-	-o	‘run’
part-	-i-	-t-	-o	‘left’	eccel-	Ø	-s-	-o	‘excelled’

Table 1: Italian past participles (based on Calabrese 2015)

The regular past participles are formed by adding the ending -to (or rather -t- and -o) to a verbal stem (highlighted in grey). The stems consist of verbal roots extended by a vowel (ThV), i.e. regular past participles are thematic verbal forms. The ThV indicates, in essence, membership of a specific conjugation class. In contrast, the irregular past participles show allomorphy in the ending (as -so vs. -to) and are athematic. The root is not extended by a ThV, which leads to shorter forms. Examples of this kind indicate a possible link between athematicity and irregularity in Romance.

In this paper, we will show that the link between athematicity and irregularity also holds for French, where the presence or rather absence of ThVs has been the subject of some debate in the literature. We will first discuss the French conjugation class system and will propose that French has two thematic and several athematic conjugation classes (CCs). As will be shown, the athematic CCs are precisely those with the (most) irregular verbs. Finally, following the DM-based Vocabulary Insertion-Only model proposed by Haugen and Siddiqi (2016), we will propose an analysis for this link between athematicity and irregularity, which we will illustrate using two well-known cases of root (or rather stem) allomorphy in French: The alternation in (2a) shows that the root final consonant of the verb is deleted in some cases (see vivre ‘to live’), but maintained in others (see arriver ‘to arrive’). In line with Schane (1966) and others, we will associate this type of allomorphy with the absence of a ThV in those cases where the root final consonant is deleted. The second example to be considered is suppletion with French aller ‘to go’, an extreme case of allomorphy. We will show that in the present tense the indicative 1st and 2nd person plural are thematic verbal forms and have the default realization for the root € ˆ (i.e. all-based forms), whereas more marked Vocabulary Items for the root € ˆ (e.g. va-based forms) are clearly athematic.

(2) (a) C/Ø-alternation: e.g. vivons [viv '@ Y V vis > Y L @ > Y Lv @
but arrivons > D' 2 L Y V arrives > D & L > D S L @
(b) Suppletion: e.g. allons [al '@ Y V vas [va @

In what follows, we will look first at the Romance conjugation classes (§2) as they are derived from Latin; we will discuss the notion of theme vowel and briefly hint at possible segmentations that have been proposed in the literature for French verbal forms (§2.1).

We will then propose our own interpretation of French CCs (§2.2), in which we consider the correlation between an assumed more complex structure with ThV-position and reduced irregular forms to be particularly insightful. In Section 3, we present a Separation approach to both the consonant-zero-alternation (§3.1) and € ˆ suppletion (§3.2). In the short conclusion (§4) we once again return to the Suppletion Generalization

2. Romance Conjugation Classes

Verbs in Latin are divided into four or five conjugation classes distinguished as a general rule by ThVs, which appear between the root and the inflectional ending (cf. Leumann [5]1977, 518ss., §398). As shown in Table 2, three CCs (I, II, IV) have a long ThV, whereas the remaining two CCs have either a short or, in some forms, an epenthetic vowel (cf. Halle 2018, Van der Spuy 2020 for more details). The third CC with /i/ is also called "consonant" or athematic, i.e. the vowel is not considered part of the stem, but is epenthetic. In addition, the thematic CCs also have athematic forms due to different phonological preferences (e.g. the avoidance of hiatus). That is, we need to distinguish between athematic CCs and athematic verbal forms.

/a:/-conjugation	/e:/-conjugation	/i/-conjugation (cons. conj., athematic)	/i:/-conjugation	mixed conjugation
cant- —re	G • E • -re	leg- -e- -re	aud- ¯-re	sap- -e- -re
cant- —mus	G • E • -mus	leg- -i- -mus	aud- ¯-mus	sap- -i- -mus
cant- À	G • E • À	leg- À	aud- -i- À	sap- -i- À

Table 2: Latin conjugation classes

Modern Romance does not preserve the historical vowel length distinction; clearly, this loss of phonemic vowel length directly affected the evolution of CCs in the Romance languages. The Latin CC system is reduced to three main classes (with subclasses, see Figure 1) in all Romance languages, albeit with different results. However, despite many diachronic and synchronic differences, Spanish, Portuguese, Catalan and Italian have

(LWKHU ¿YH &&V LQFOXGLQJ WKH VR FDOOHG PL[HG F WKH WKLUG DQG WKH IRXUWK && RU IRXU H[FOXGLQJ L the third and fourth class into one, hence assuming only 3 CCs. If the division were based on the presence or absence of ThV, i.e. thematic and athematic CCs, we would have only 2 CCs (cf. Kühner and Holzweissig 1912, 659). Note, furthermore, that the notion "thematic" is not used in a uniform way in linguistics.

clearly preserved the thematic conjugation system.[5] Additionally, in all these languages the respective CCs are not equally productive or equally regular and they also differ in their number of members. The first CC, i.e. that marked by ThV -a-, is the less marked in all these languages. It is the productive CC, it has (nearly) no irregular verbs and it has the largest number of members. Moreover, the opposition between the other CCs may be neutralized in some tenses (e.g. Sp. parl*ó* and *temía*). Based on these observations, Oltra Massuet (1999, 2000) argues that ThVs are bundles of subatomic abstract features organized in a markedness hierarchy or feature geometry (see Figure 1; cf. Pomino and 5HPEHUJHU >VXEPLWWHG@ IRU PRUH GHWDLOV

Figure 1: Conjugation Classes (regular verbs): Spanish (a), Catalan (b) (Oltra Massuet 2000, 287) and Italian (c) (based on Napoli and Vogel 1990)

ThVs in French have developed differently. To date, there has been no consensus in the literature with respect to whether or not Modern French has ThVs or how the CCs are to be described. In what follows, we will present our line of reasoning with regard to these two questions without discussing alternative analyses in detail for reasons of space FI 3RPLQR DQG 5HPEHUJHU >VXEPLWWHG@ IRU D GH

2.1 Are There ThVs in French?

There are two main proposals in the literature with regard to possible segmentations RI WKH LQILQLWLYDO IRUPV RI)UHQFK YHUEV VHH 7 discussion): The "outcome" of the former Latin ThVs can be (a) attributed to the verb ending (see proposal II) (e.g. Michaut 1934; Dupuis 1935; Larousse 1936; Grevisse 1969; Schane 1968; Paradis and El Fenne 1992), (b) added to the root to form a stem

5RPDQLDQ KDV QRW RQO\ SUHVHUYHG ¿WKH &RXDQ&V YDEH[DXJPHQWHG VXEFODVVHV FI 3DQ 'LQGHOHJDQ

(i.e. root+ ThV; see proposal I) or (c) seen as part of a non-divisible stem (i.e. e.g. aimer > - P r @ H J 0 D U W L Q H W 9 D Q G H Q (\ Q G H D Q G % O D Q Coute 1981; Plénat 1981). That is, some linguists follow the rather traditional classification and consider -r to be the infinitival ending, with the preceding vowel classified as the theme vowel/theme diphthong, while others propose that -er, -ir, -oir and other endings should be analyzed as purely infinitival, i.e. inflectional.

Latin	French	proposal I	proposal II	general rule (not exceptionless)
D P — U H a i m e r	> - P r @	or > - P r @ H		Latin -—- developed to -
V H Q W ¯ s e n t i r	> V $ W L U @	> V $ W L U @		Latin - ¯- “remained”--
Y D O • U H v a l o i r	> Y D O Z D U @	or > Y D O Z D U @		- • G L S K W K R Q J L] H V W R (“thematic diphthong”)
scrib-e-re >	écrire	> H N U L U @		-e- (or rather /i/) is “lost” (= athematic)

Table 3: From Latin to French (cf. a.o. Foley 1979, 135)

In recent years, some doubt has been cast on the idea that French has TVs; one reason for this is that finite forms are at least superficially athematic (see Table 4). One question to be answered, however, is whether we are dealing with athematic forms or complete athematic conjugations. As illustrated above, Latin thematic CCs also have athematic forms.

	aimer	sentir	valoir	écrire
1sg	aime > - P @	sens > V $ @	vaux [vo- z @	écris [ekri- z @
1pl	aimons > - P z @	sentons > V $ W @	valons > Y D z @	écrivons > H N U L Y '

Table 4: Selected finite verbal forms

One source of evidence for the presence of ThVs in the first CC in French comes from the consonant-zero-alternation which we will analyze in Section 3.1: The root final consonant of athematic viv(re) ‘to live’ (3rd CC) is maintained if there is a possibility for it to appear in a syllable onset, i.e. before V (e.g. nous vivons > Y L Y z @ Y L Y H O L Y H ¶

% R \ p D V V X P H V V H S D U D W H V W H P V I R U L Q ¿ Q L W L Y D C W R N H H S L Q À H F W L R Q D O W K H G L Q ¿ J Y L W L J Y H D V U H J X O D U D V S L Q U H J X O D U Y H U E V G H U L Y H V I U R P ³ W K H X Q L ¿ F D W L R Q R I I U H S U H V H Q W D W L R Q V E U R X J K W I R U Z D U G E \ W K H G L V W L Q F W L D Q G À R D W L Q J H Q W L W L H V ´ R X U W U D Q V O D W L R Q) R U % play a role in the slots of their “stem space”. Starke (2020), in a nanosyntactic approach, also has no representations for ThVs.

before a C with which it can build a complex onset (e.g. nous vivrons > Y L YLYRQ ‘we live’), but is deleted before a following consonant with which it cannot form an onset (tu vis > Y L] sg @ O L \ Y ¶ Q R W tu vivs > Y L Y] @ 7 K H V D P H I L Q arriv(er) (1st CC), in contrast, is never deleted: tu arrives > D S L Y sg arrive @ not \ R X tu arris > D S L] @ 6 F K D Q H D Q G R W K H U V D V V X P H W G R H V Q R W V X U I D F H E O R F N V F R Q ¥ R Q D Q W G H O H W L R Q

The assumption that the 2nd CC (e.g. finir) contains a ThV is straightforward, since K H U H W K H 7 K 9 V X U I D F H : L W K W K H W K H F R Q ¥ R Q D Q W > W @ R U K H L V L @ follow Schwarze (2009), who assumes that the underlying form of the ThV is /is/ in all F D V H V 7 K H V R I W K L V 7 K 9 s Z K L F K L V D O V R V R P H D X J P H Q W s V X U I D F H V R Q O \ Z K H Q L W F D Q R F F X S \ D (see the plural forms of finir in Table 5), otherwise it is deleted.

	morphological structure			syllable structure	surface form[6]		spelling
	root	ThV	3		no liaison	liaison	
1sg	fin	is	z	fi.nisz	> I L Q L @	> I L Q L]	finis
2sg	fin	is	z	fi.nisz	> I L Q L @	> I L Q L]	finis
3sg	fin	is	t	fi.nist	> I L Q L @	> I L Q L W	finit
1pl	fin	is	'	fi.ni.sonz	> I L Q L @	> I L Q] L 9 @	finissons
2pl	fin	is	ez	fi.ni.sez	> I L Q L ¥ H	> I L Q L W H] H @	finissez
3pl	fin	is	% W	I L Q L V ¥ W	> I L Q L V @	> I L Q L V W	finissent

Table 5: Forms of finir (adapted from Schwarze 2009)

---

7 K H U H D U H D O W H U Q D W L Y H D Q D O \ V H V I R U W K L V N L Q G R stems in El Fenne (1994), stem spaces in Bonami and Boyé (2002) and in Bonami et al. (2008) and consonant epenthesis in Klausenburger (1974), Tranel (1974, 1976), Kaye and Morin (1978), Morin and Kaye (1982) among others. Bonami and Boyé (2002), for example, derive stem dependencies from conjugation patterns, which roughly correspond to what are called P R U S K R P H V F I \$ U R Q R u R U P R U S K R P L F S D W W H U Q V approaches to morphology. We cannot discuss further details of these approaches here for reasons R I V S D F H E X W V H H 3 R P L Q R D Q G 5 H P E H U J H U > V X E P L W W H G @

5 Again, not all linguists assume ThV for the second conjugation (see Fn. 3) (e.g. El Fenne 1994; Bonami and Boyé 2002; Bonami et al. 2008).

6 In the phonic (= spoken, as opposed to the graphic/written modality) realization of French, the phenomenon of liaison is one of the most striking sandhi phenomena of the language. Liaison is X Q G H U V R R G D V W K H D W H Q W Z R U G L ¿ Q D O R F F L S D Q W Z K L F K prosodic context) is not pronounced before a following word-initial consonant, but is realized before a following word-initial vowel.

Interestingly, the ThV of the second conjugation class spreads across other CCs in certain tenses, leading to a neutralization of the CC opposition (see Figure 2). The first conjugation class, however, is never affected by this neutralization process. This means that in some tenses we have a twofold opposition (i.e. 1st CC vs. other CC) and in others a threefold opposition (i.e. 1st CC vs. 2nd CC vs. other CC) and so on.

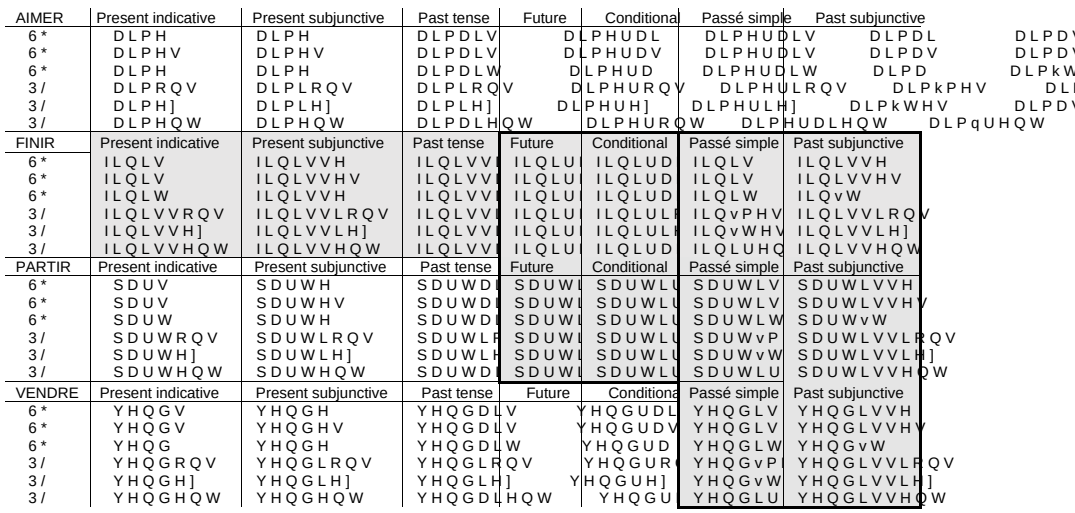


Figure 2: CC oppositions and neutralizations[7]

2.2 The French Conjugation Class System

Considering the paradigms in Figure 2, we assume that French CCs are also hierarchically structured according to diacritic features. As shown by Oltra Massuet (1999) for Catalan, the respective features allow us to understand the neutralizations of the CCs mentioned previously and to grasp differences in markedness (i.e. productivity, irregularity, size). We propose the hierarchy in Figure 3, which is not restricted to regular verbs, but also includes irregular verbs.

7 We are aware of the fact that the passé simple and the past subjunctive are archaic and no longer used in Modern French. Nevertheless, this neutralization helps us to understand how the CC system is structured in French.

Figure 3: Feature geometry for the French conjugation classes

According to this specification, CC neutralization in the passé simple and past subjunctive U H O L H V R Q W K H F R Q W U D V W L Q L W L D O V W H P V R I finir, partir and vendre, whereas the neutralization L Q W K H I X W X U H D Q G F R Q G L W L R Q D O P R R G R I aimer and aller to W K H

At the same time, the proposed feature geometry also captures different degrees of markedness of the respective CC (from less marked to more marked; see Figure 4): aimer > finir > partir > vendre > aller.[8] This allows us to understand the link between athematicity and irregularity: The most marked CCs contain the (most) irregular verbs and it is precisely these CCs, leaving neutralizations aside, that are athematic. Put another way, thematic aimer and thematic finir belong to those CCs whose verbs are considered either fully regular or "irregularly" regular, i.e. verbs that undergo minor (morpho) phonologically predictable changes (e.g. vowel alternations: lever > O Ñ Y H @ µ W R U D lèvent > O - Y @ µ W R U). The athematic verbs partir and vendre belong instead to CCs that contain irregular verbs showing other subregularities. Finally, suppletive aller (with a mixed paradigm of forms from thematic and athematic verbs) belongs to the residual CC, which is a ragbag of irregular verbs with idiosyncratic alternations (cf. Gertner 1973 and Meunier and Marslen-Wilson 2004 for a classification of French irregular verbs).

8 The degree of markedness results from the total number of features and their respective values (+ or -). For instance, partir K D V W Z R > @ I H D W X U H V D Q G W Z R > @ I P D U N H G Z L W K > @ L Q G H [K D V R Q Z K I H D W X U H D Q G W Z R > R I D U H P D U N H G Z L W K > @ L Q G H [W K U H H D Q G @ I H D W X U H V L H R X W R I L V P D U N H G Z L W K > @ L Q G H [7 K L V D O W H U Q D W L R Q F I Semerdjian-Vaissière (2002) on vowel lowering in closed syllables (Schwarze and Lahiri 1998).

Figure 4: Markedness of CC, athematicity and irregularity

We assume that neither the syntactic derivation nor the post-syntactic well-formedness conditions are affected by the conjugation class features. This means that all CCs have the same morphological structure for Vocabulary Insertion (VI) (see Figure 5).

Figure 5: Morphological structure

The realization of this structure differs considerably between CCs, however. As Table 6 shows, the 1st and 2nd CCs (atomis*er*, *ir*, *finir*) have a regular root followed by three inflectional affixes in the 1st plural imperfect. The CCs that are marked by two diacritics, instead, allow fewer affixes after the “irregular” root, i.e. the ThV is missing. What is more, as exemplified by the present tense forms of the 2nd person, only these athematic verbs show irregular forms with subregularities (*partir* and *vendre*; also *peindre*) or with idiosyncrasies (as in the case of suppletion with *aller*).

			morphological structure					syllable structure	surface form	
			root	v°	ThV	T°	3		no liaison[10]	spelling
	>	. @	atom	is	%	i	' ½ z	D W R P ½ L	> V D % W L R ½ @	atomisons
	>	. @	- P		%	i	' ½ z	- P % ½ L '	> - P ½ @ M '	aimions
1pl	>	.	fin	/ @	is	i	' ½ z	I L Q ½ L V L	> ' I L Q ½ @ L V	finissons
impf.	>	.	S	/ @ S W		i	' ½ z	S D S ½ W L '	> S D S ½ @ W M	partions
	>	.	@ Y $	G		i	' ½ z	Y $ ½ G ½ '	> Y z $ ½ @ M '	vendions
	>	. @	al		%	i	' ½ z	D O ½ % L '	> D ½ @ O M '	allions
	>	. @	atom	is	%		' ½ z	D W R P ½ L	> V D % W R ½ @	atomisons
	>	. @	- P		%		' ½ z	- P % ½ '	> - P ½ @	aimons
1pl	>	.	fin	/ @	is		' ½ z	I L Q ½ L V '	> I L Q ½ @ L V	finissons
pres.	>	.	S	/ @ S W			' ½ z	S D S ½ W '	> S D S ½ @ W	partons
	>	.	@ Y $	G			' ½ z	Y $ ½ G ½ '	> Y z $ ½ @	vendons
	>	. @	al		%		' ½ z	D O ½ % '	> D ½ @ O '	allons
	>	. @	atom	is	%		z	D W R P L	> V D % W R	atomises
	>	. @	- P		%		z	- P %]	> - P @	aimes
2sg	>	.	fin	/ @	is		z	fi.nisz	> I L Q L @	finis
pres.	>	.	S	/ @ S W			z	S D S W]	> S D S @	pars
	>	.	@ Y $	G			z	Y $ ½ G ½]	> Y ½ @	vends
	>	. @	va				z	vaz	> Y D @	vas

Table 6: Relation between morphological structures and the CC

This means that Vanden Wyngaerd's generalization, according to which irregularity in form of either the root or the suffix goes hand in hand with a reduced number of suffixes, is also valid for French verbal inflection.

3. Non-Terminal Vocabulary Insertion

In this section, we will discuss how the generalization made so far can be captured in the framework of Distributed Morphology (DM; Halle and Marantz 1993, 1994). Based on data from suppletion in particular, we follow the assumption that functional elements and roots are both inserted post-syntactically (cf. Haugen and Siddiqi 2013). As is well-known, in the standard version of DM a set of post-syntactic operations can alter the syntactic output before VI, including Fusion, Impoverishment, and Pruning,

10 We have omitted the liaison forms for reasons of space.

11 The denominal verb atomiser 'to atomise', which has no root augment, but a verbalizer inserted in v°. This derivation again, belongs to the productive CC 1 and therefore comes with a ThV.

among others. For the athematic Italian past participles mentioned above, such as corso, Calabrese (2015) proposes an analysis based on Pruning. Apart from the fact that the arguments in Calabrese (2015) are somehow circular,[12] his analysis is not applicable to French ‹ suppletion. However, other post-syntactic processes, such as Fusion, for example, also seem incorrect for the analysis of French suppletion, since locality restrictions on allomorphy are not met (but see Pomino and Remberger 2019). What is more, Fusion and Pruning can be avoided if VI is not limited to terminal elements, as in the non-terminal spell-out of Nanosyntax (cf. Starke 2009) and the Spanning approach of Svenonius (2012), among others. In sum, the assumption of multiple post-syntactic processes has met with some criticism and, given equal explanatory adequacy, a smaller number of processes is preferable to a larger number (cf. Haugen and Siddiqi 2016, 346; Trommer 1999 a.o.). Haugen and Siddiqi (2016) argue that the post-syntactic processes of DM should be reduced to VI only. Following Svenonius (2012, 2016) as well as Haugen and Siddiqi (2016), we will integrate Non-Terminal Insertion or rather Spanning[13] into the DM framework in order to explain the consonant-zero-alternation and suppletion in French.

3.1 Spanning and the Consonant-Zero-Alternation in French

Spanning is a specific type of VI. As in the standard version of DM, this mapping process operates over the hierarchical structure, but it allows the insertion of phonological material not only in one terminal node at a time but also in spans of terminal nodes that are in a complement relation with each other (Williams 2003; Svenonius 2012; Merchant 2015). For ease of illustration, let us assume that the output of the syntactic derivation is the complex head given in Figure 6. In the Spanning approach, VI can target any of the listed spans, i.e. the root may be realized alone or together with X (or X and Y etc.).

12 Pruning, according to Calabrese (2015), takes place before VI, i.e. it cannot be conditioned by the context of VI; this raises the question of why the input structure for VI is reduced only in particular (categorial or non-categorial) contexts.

13 For French morphology in Nanosyntax, cf. Starke (2020). Starke (2020) does not seem to posit CCs, nor does he consider ThVs or latent consonants in his structures. We do not adopt a nanosyntactic approach since we assume that morphophonological exponence cannot be directly represented in syntactic structures, but it is syntax that serves as an input to VI, which is guided by the Subset Principle. Nevertheless, some of the ideas from Nanosyntax, e.g. the insertion of UHJXODU DQG LUUHJXODU OH[LFDO LWHPV LQWR VHYHUDO QRGHV DW RQFH DSSURDFK WR VRPH H[WHQW DOWKRXJK WKH SURSRVDO WKDW ZRUGV UHSUHVHQW QRQ6WDUNH OH[LFDOL]H OHVV VWUXFWXUH LQ 6WDUNH

Figure 6: Spanning

We will assume that the difference between regular and irregular verbs in French (and other languages) depends on whether or not the output structure contains a one-to-one realization of terminal nodes. More precisely, as soon as Root and X° are realized as a span and not as two separate slots, the verbal forms lose their structural transparency and may be irregular. Our analysis is based on two central assumptions: (i) Whether Vocabulary Items target only the Root or span over the Root and ThV depends on the CC features H J > . @ H Q F R G H G L Q W K H U R R W L L ' H S H Q G L Q J R[14] This idea is exemplified in Figure 7: In the case of aimer/arriver and finir, VI for the root realizes only the Root and the root-final consonants surface, since they always precede a ThV. In contrast, in the case of partir/vendre, VI spans over Root and Th and the root-final consonants may not surface. More precisely, the consonant is never saved by a ThV (which offers an open onset position) and its (non)appearance depends on K R Z W K H 7 3 L V U H D O L] H G , W Z L O O V X U I D F H R Q O \ Z L position that the consonant could move into (e.g. -ons and -ez, but also -r-ons and -r-ez) and will otherwise be deleted.

14 This is true at least for the present tense, which can also be interpreted as a non-tense and therefore usually has no visible exponent.

Figure 7: Spanning and the consonant-zero-alternation

With this analysis we can capture the link between athematicity and irregularity in the following way: ThVs prevent the deletion of root-final consonants. As soon as VI spans over — ‹ ˆ ˆ •+v°+ThV, the realization of the final consonant depends on the realization of the following slot (or span) and, since this slot does not per definition begin with a vowel, the final consonant may be deleted leading to “irregular” verbal forms.

3.2 Spanning and French GO Suppletion

The suppletion found with French aller is especially interesting (cf. also Pomino and Remberger 2019) since the respective forms have their origins in verbs from different Latin CCs, i.e. some forms of aller are thematic and others are athematic. The three source verbs are: Y — G (athematic), Y — G (athematic) and D P E X O — U H D P E L W — U H D Q G — U H (thematic). In this suppletive pattern, the link between athematicity and irregularity would thus be expected to be particularly transparent.

In French, suppletion in the verbal forms of is, at least at first glance, sensitive to person and number (but only in the present tense and in the imperative, i.e. non-categorical suppletion; cf. Veselinova 2006) and is additionally sensitive to TAM (i.e. categorical suppletion; cf. Veselinova 2006): Forms stemming from Latin , ‹ ~ are attested in French only in the future and the conditional, while those originating from athematic Y — G are found in the present indicative (and in the imperative), but only in the singular and 3pl. All other forms are based on thematic D O O — U H which we consider to be the default realization for the root in French. In the following analysis, we will look more closely at the present tense forms, since here we have both athematic and thematic suppletive stem-allomorphs, which seem to be contextually sensitive to 3 I H D W X U H V V H H 7 D E O H

	present ind.		present subj.		imperfect		future	
	sg	pl	sg	pl	sg	pl	sg	pl
1	vais	all-ons	aille	all-i-ons	all-ais	all-i-ons	i-r-ai	i-r-ons
2	vas	all-ez	ailles	all-i-ez	all-ais	all-i-ez	i-r-as	i-r-ez
3	va	vont	aille	aillent	all-ait	all-aient	i-r-a	i-r-ont

Table 7: Selected verbal forms of suppletive aller (athematic forms highlighted in grey)

It is very interesting to note that the present tense forms of French € ˆ in particular have a very reduced segmentability, i.e. there is a considerable mismatch between the hierarchical structure (e.g. — ‹ ˆ ˆ • Y ƒ 7 K 9 7 3 V H H)- L J X U H ophonological realization of this structure. More precisely, the all-based forms have a higher degree of segmentability than the va-based forms and this, again, goes hand in hand with (a)thematicity: The root all- selects ThV, whereas va- is athematic. In the Spanning approach, this means that the root all- realizes only the — ‹ ˆ ˆ • slots, while va- spans over — ‹ âˆ• ThV; see Figure 8.

Figure 8: Spanning in case of suppletion

In the case of non-categorial suppletion, as in the present tense of aller, it may also be the context of the agreement features that conditions the insertion of suppletive VWHPV 6LQFH WKH 7ƒ DQG 3 DUH D VSDQ LQ WKH SU the adjacent span. In the more specific insertion context in the singular and 3rd person plural more specific Vocabulary Items are inserted. In contrast, the 1st and 2nd plural UHSUHVHQW WKH GHIDXOW L H 9, aller belongs to the QRW FRQ WH[W WKHPDWLF && LQFOXGLQJ WKH GLDFULWLF > .@ WKH SKRQRORJLFDO UHDVRQV ZKLFK PHDQV WKDW 3 LV Q not influence VI of the root.

4. Conclusion

We have argued that French has two thematic and several athematic conjugation classes which are hierarchically structured. Based on this, we have additionally shown that the Suppletion Generalization (Vanden Wyngaerd 2018, 1) also holds for French: “If there is irregularity in the form of either the root or the suffixes, the number of suffixes gets reduced.” This generalization can be integrated into the DM framework if we admit non-terminal insertion or rather Spanning. We have shown how athematicity of the French CC – which results from VI spanning over — ‹ ˆ ˆ • and ThV – impinges on the irregularity of the respective forms. (Ir)regularity can be explained by different spanning sizes of the respective roots, which is dependent on the CC-features of the roots.

5. Works Cited

Aronoff, Mark. 1994. Morphology by Itself. Cambridge, MA: MIT Press.

Bobaljik, Jonathan D. 2000. “The Ins and Outs of Contextual Allomorphy.” In University of Maryland Working Papers in Linguistics 10, edited by Kleanthes Grohmann and Caro Struijke, 35–71. Maryland: Linguistics Dep., University of Maryland.

Bobaljik, Jonathan D. 2012. Universals in Comparative Morphology. Cambridge, MA: MIT Press.

Bonami, Olivier, and Gilles Boyé. 2002. “Suppletion and Stem Dependency in Inflectional Morphology.” In Proceedings of the HPSG ‘01 Conference, edited by Franck Van Eynde, Lars Hellan, and Dorothee Beerman, 51–70. Stanford: CSLI Publications.

Bonami, Olivier, Gilles Boyé, Hélène Giraudo, and Madeleine Voga. 2008. “Quels verbes sont réguliers en français?” Congrès Mondial de Linguistique Française 2008, edited by Durand Jacques, Benoît Habert, and Bernard Laks. Paris: Institut de Linguistique Française. Accessed October 29, 2021. https://www.linguistiquefrancaise.org/articles/cmlf/pdf/2008/01/cmlf08186.pdf

Boyé, Gilles. 2000. 3 U R E O q P H V G H P R U S K R S K R Q R O R J L H Y H U I italien. Ph.D. thesis, Université Paris 7.

Caha, Pavel, Karen De Clercq, and Guido Vanden Wyngaerd. 2019. “The Fine Structure of the Comparative.” Studia Linguistica 73 (3): 470–521.

Calabrese, Andrea 2015. “Irregular Morphology and Athematic Verbs in Italo-Romance.” , V R J O R V V 6 S H F L D O , V V X H R Q 69 - W 02. O R 5 R P D Q F H
Dupuis, D. 1935. La grammaire pour tous. Paris: J. de Gigord.
El Fenne, Fatimazohra. 1994. La flexion verbale en français: contraintes et stratégies de réparation dans le traitement des consonnes latentes. Ph.D. diss., Université Laval.
Foley, James. 1979. Theoretical Morphology of the French Verb. Amsterdam: John Benjamins.
Gertner, Michael H. 1973. The Morphology of the Modern French Verb. The Hague: Mouton.
Grevisse, Maurice. 1969. Précis de grammaire Française. 28th ed. Duculot, Gembloux.
Halle, Morris, and Alec Marantz. 1993. “Distributed Morphology and the Pieces of Inflection.” In The View of Building 20: Essays in Honor of Sylvain Bromberger, edited by Kenneth Hale and Samuel J. Keyser, 111–176. Cambridge, MA: The MIT Press.
Halle, Morris, and Alec Marantz. 1994. “Some Key Features of Distributed Morphology.” MIT Working Papers in Linguistics: Papers on Phonology and Morphology 21, 275–288.
Halle, Morris. 2018. “Aspects of the Morphophonology of the Verb in Latin and in German and English.” Linguistic Inquiry 50 (1): 3–12.
Haugen, Jason D., and Daniel Siddiqi. 2013. “Roots and the Derivation.” Linguistic Inquiry 44 (3): 493–517.
Haugen, Jason D., and Daniel Siddiqi. 2016. “Towards a Restricted Realization Theory. Multimorphemic Monolistemicity, Portmanteaux, and Post-Linearization Spanning.” In Morphological Metatheory, edited by Daniel Siddiqi and Heidi Harley, 343–386. Amsterdam/Philadelphia: Benjamins.
Kaye, Jonathan, and Yves Charles Morin. 1978. “Il n’y a pas de règles de troncation, voyons!” In Proceedings of the Twelfth International Congress of Linguists, edited by Wolfgang Dressler and Wolfgang Meids, 788–792. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
Klausenburger, Jürgen. 1974. “Rule Inversion, Opacity, Conspiracies: French Liaison and Elision.” Lingua 34: 167–179.
Kühner, Raphael, and Friedrich Holzweissig. 1912 (2021). Ausführliche Grammatik der lateinischen Sprache. Darmstadt: wbg.
Larousse. 1936. Grammaire Larousse du français contemporain. Paris: Larousse.
Leumann, Manu, Johann Baptist Hofmann, and Anton Szantyr. 1977. Lateinische Grammatik. Erster Band: Lateinische Laut- und Formenlehre. München: C. H. Beck.
Maiden, Martin. 2016. “Morphomes.” In 7 K H 2 [I R U G * X L G H W R W K H 5 R P D Q F H by Martin Maiden and Adam Ledgeway, 708–721. Oxford: Oxford University Press.
Martinet, André. 1969. Le français sans fard. Paris: Presses Universitaires de France.
Merchant, Jason. 2015. “How Much Context Is Enough? Two Cases of Span-Conditioned Stem Allomorphy.” Linguistic Inquiry 46 (2): 273–303.

Meunier, Fanny, and William Marslen-Wilson. 2004. “Regularity and Irregularity in French Verbal Inflection.” Language and Cognitive Processes 19 (4): 561–580.
Michaut, Gustave. 1934. Grammaire française. Paris: Hatier.
Morin, Yves Charles, and Jonathan Kaye. 1982. “The Syntactic Bases for French Liaison.” Journal of linguistics 18: 291–330.
Napoli, Donna Jo, and Irene Vogel. 1990. “The Conjugations of Italian.” Italica 67: 479–502.
Oltra Massuet, Isabel. 2000. “Conjugation Class.” In [illegible] Linguistics. Oxford: Oxford University Press.
Oltra Massuet, Maria Isabel. 1999. “On the Notion of Theme Vowel: A New Approach to Catalan Verbal Morphology.” MA Thesis. MIT, Cambridge, MA.
[illegible] sity Press.
Paradis, Carole, and Fatimazohra El Fenne. 1992. “L’alternance C/Ø des verbes français: une analyse par contraintes et stratégies de réparation.” [illegible] tique 21: 107–141.
Paradis, Carole, and Fatimazohra El Fenne. 1995. “French Verbal Inflection Revisited: Constraints, Repairs and Floating Consonants.” Lingua 95: 169–204.
Pinchon, Jacqueline, and Bernard Coute. 1981. Le système verbal du français: description et application pédagogique. Paris: Nathan-Université.
Plénat, Marc. 1981. “L’ ‘autre’ conjugaison ou de la régularité des verbes irréguliers.” Cahiers de grammaire 3: 1–304.
Pomino, Natascha, and Eva-Maria Remberger. Submitted. “Does French have Theme Vowels?” Ms. to be published in Theme Vowels, edited by Svitlana Antonyuk [illegible]
Pomino, Natascha, and Eva-Maria Remberger. 2019. “Verbal Suppletion in Romance Diachrony: The Perspective of Distributed Morphology.” Life-Cycle of Suppletion. Special Issue of the Transactions of the Philological Society 117 (3), edited by Frans Plank, and Nigel Vincent, 471–497.
Schane, Sanford A. 1966. “The Morphophonemics of the French Verb.” Language 42 (4): 746–758.
Schane, Sanford A. 1968. French Phonology and Morphology. Cambridge, MA: MIT Press.
Schwarze, Christoph, and Aditi Lahiri. 1998. Einführung in die französische Phonologie. Arbeitspapier Nr. 88. Konstanz: Fachgruppe Sprachwissenschaft der Universität Konstanz.
Schwarze, Christoph. 2009. “The French i-Conjugation from a Diachronic Perspective.” In Selected Proceedings of the 6th Décembrettes, edited by Fabio Montermini, Gilles Boyé, and Jesse Tseng, 35–49. Somerville, MA: Cascadilla Proceedings Project.

Starke, Michal. 2009. “Nanosyntax: A Short Primer to a New Approach to Language.” Nordlyd 36: 1–6.
Starke, Michal. 2020. “UM – Universal Morphology.” Keynote talk at NELS 51, 6.-8.11.2020, University of Québec in Montréal (UQAM).
Svenonius, Peter. 2012. “Spanning.” Ms., CASTL, University of Tromsø.
Svenonius, Peter. 2016. “Spans and Words.” In Morphological Metatheory, edited by David Siddiqi and Heidi Harley, 201–222. Amsterdam: John Benjamins.
Tranel, Bernard. 1974. The Phonology of Nasal Vowels in Modern French. PhD diss. University of California. San Diego.
Tranel, Bernard. 1976. “A Note on Final Consonant Deletion, the Pronunciation of Cardinal Numbers, and Linguistic Change in Progress in Modem French.” Lingua 39: 53–68.
Trommer, Jochen. 1999. “Morphology Consuming Syntax’ Resources: Generation and Parsing in a Minimalist Version of Distributed Morphology.” Accessed October 14, 2021. https://ling.auf.net/lingbuzz/000116.
Van den Eynde, Karel, and Claire Blanche-Benveniste, 1970. “Essai d’analyse de la morphologie du verbe français. Présentation d’hypothèses de travail.” Orbis XIX (2): 404–429.
Van der Spuy, Andrew. 2020. “The Theme Vowels of Latin verbs.” Lingua 243. Accessed June 2022. https://doi.org/10.1016/j.lingua.2020.102883.
Vanden Wyngaerd, Guido. 2018. “Suppletion and Affix Reduction.” Accessed October 29, 2021. https://lingbuzz.net/lingbuzz/004095.
Veselinova, Ljuba. 2006. Suppletion in Verb Paradigms: Bits and Pieces of a Puzzle. Amsterdam: John Benjamins.
Williams, Edwin. 2003. 5 H S U H V H Q W. Cambridge, MA: MIT Press.

Ž YV D $OWHUQDWLRQ LQ WKH $X[DQG 'LVWULEXWLRQ RI 2&OV LQ &D 'LDOHFWV RI /DXVEHUJ $UHD

/HRQDUGR 0 D6DYRLD [a], %HQHGHWWD %DOGL [b]

University of Florence, Italy

[a]leonardomaria.savoia@unifi.it, [b]benedetta.baldi@unifi.it

Abstract. This article deals with the realization of 3rd person object clitics in auxiliary contexts in some Southern-Italian dialects spoken in “Lausberg area”, where the DOORPRUSK\ - D LQ WKH DX[LOLDU\ VWHP DQG HQ phenomenon in relation to the overall distribution of OCIs. A crucial question is the theoretical status of morphology. Although morphology is nothing more than a way of expressing syntactic structures, it is traditionally seen as a syntactic component, that, according to DM, conveys an information “separated from the original locus of that information in the phrase marker” (Embick and Noyer 2001, 557) by means of rules PDQLSXODWLQJ V\QWDFWLF QRGHV :H SURSRVH D GLI governed by the same computational rules of syntax and subword elements are fully interpretable. Inflection and clitics are the result of (pair-)merge operations (Chomsky D E JLYLQJ ULVH WR DPDOJDPV EDVHG RQ DJUH

Keywords: Object clitics; DOM; morphology; auxiliary; merger operation; Lausberg area varieties

,QWURGXFWLRQ

The topic of this article is the realization of object clitics in auxiliary contexts in Southern-Italian dialects spoken in Lausberg Area. The issue at stake is the nature of morphology and its theoretical status. This is a long-term question, dating back at least to Chomsky (1972), when the lexicalist hypothesis became the general solution in the treatment of the relation between syntax and interpretation. The portion of morphology that remained within syntax however has kept representing a problem for the analysis. In fact, morphology is necessary for syntax but it seems to apply specific structural

requirements, which according to some authors respond to a linear adjacency criterion (Embick and Noyer 2001) rather than to structural organization. Moreover, morphology shows a set of phenomena concerning the order of exponents, syncretism and other properties apparently idiosyncratic, not immediately associated with functional or structural codification. In the following discussion, we will apply an approach that holds to the assumption that morphology is part of the linguistic knowledge, as it is governed by the same fundamental computational tools of syntax.

The diverse morpho-syntactic realizations of 3rd person OCls in the Lausberg area dialects (cf. Lausberg 1939) involve the alternation between internal inflection and clitic, providing us with a test bench as regards the ability of the syntactic procedures to construct morphologically complex words. The article, after presenting the data and describing the different systems, discusses some theoretical points in order to clarify the salient aspects of the model we adopt, and finally it proposes an analysis of these phenomena. The aim is to reach a unified and general treatment.[1]

&OLWLFV DQG ,QWHUQDO ,QIOHFWLRQ RI WKH

In Romance dialects spoken in villages within the Lausberg area, on the border between Calabria (Morano and Albidona) and Basilicata (Colobraro), 3rd person internal arguments (IA) show a specialized morphological mapping in auxiliary contexts. In some varieties, in the active form of transitive verbs the auxiliary have alternates two allomorphs, one with the stem vowel a- DQG RQH ZLWK - incorporates the realization DQW of the 3rd person, as in (1a, b, c) for Morano. In the examples, we indicate with the ‰ Œ exponent (or referent) of 3rd person, with Œ € Œ € ‰ … ‰ … the verbal agreement e or the OCls. , ‡ is the gloss of the element © that realizes the final vocalic exponents in many of these dialects; † Œ € • Œ € ‰ … correspond to the agreement nominal fe

(1) (a) D 1 ‡ DXP - W X D L
(3 ‰ Œ) have-1 called-† Œ € • Œ € ‰ …
µ , KDYH FDOOHG KLP KHU WKHP ¶

(b) a: FDP - W X D L
(3 ‰ Œ) have called-† Œ € • Œ € ‰ …
µ \RX KDYH FDOOHG KLP KHU WKHP ¶

---

[1] 7KH DUWLFOH LV WKH IUXLW RI FRPPRQ UHÀHFWLRQ DQ ZRUN KDYH EHHQ FROOHFWHG WKURXJK ¿HOG LQYHVWLJDW sincere gratitude.

[2] :H XVH EHU DG SKRQHWLF WUDQVFULSWLRQ LQ JHQHUDO syllables are long; a variable lengthening may characterize also pre-tonic vowels.

(c) a: F F D P - W X D L
(3 ‰ Œ) have called-† Œ € • Œ € ‰ …
µ V K H K D V F D O O H G K L P K H U W K H P ¶ Morano

: K L O H L Q W K H G L D O H F W R I 0 R U D Q R W K L V L V W K H R Q
these contexts, in other dialects we find two ways to introduce the object. Thus, in the dialect of Colobraro in (2), D S allomorphy characterizes the 3rd singular person of the auxiliary, in (2a), and is in complementary distribution with the realization of 3rd person
2 & O V L Q H Q F O L V L V R Q W K H V W Q G D Q G V W D Q G U
The 2nd plural in turn excludes the realization of the OCl, however presenting the only alternant with a-, as in (2c)

(2) (a) a: F F D P - W %
(3 ‰ Œ) have.3 Œ € called , ‡ • …
µ V K H K D V F D O O H G K L P K H U W K H P ¶

(b) - G G d - M X - P - Q D F D P - W %
have.1 Œ € Œ € ‰ .†. Œ € ‰ •. Œ € called-, ‰ • ….
µ , K D Y H \ R X K D Y H Z H K D Y H W K H \ K D Y H F D O O

(c) D Y H V % F D P - W %
(3 ‰ Œ) have.2 ‰ … called-, ‡ • …
µ \ R X K D Y H F D O O H G K L P K H U W K H P ¶ Colobraro

7 K H D O W H U Q D Q W - R F F X U V L Q D O O R W K H U F R Q W H [
X Q H U J D W L Y H V D V L Q D D ¶ D Q G E E ¶

(3) (a) - 1 1 X vinut-u
have-1 Œ € come-† Œ €
µ , K D Y H F R P H ¶

D ¶ 1 1 X rurmut-u
have-1 Œ € slept- † Œ €
µ , K D Y H V O H S W ¶ Morano

(b) - Y Y % Q X W %
have.3 Œ € come-, ‡ • …
µ V K H K D V F R P H ¶

E ¶ - G G X U P X W %
have.3 Œ € slept- , ‡ • …
µ V K H K D V V O H S W ¶ Colobraro

0 R U H R Y H U - F K D U D F W H U L] H V D [illegible] O H [L F D O ' 3 V R F F X U L Q E ¶

(4) (a) t - 1 1 X vist-u
2 Œ € have-1 Œ € seen- † Œ €
µ , K D Y H V H H Q \ R X ¶

(b) - 1 1 X vist-u a u U D W W W %
Have-1 Œ € seen- † Œ € to brother-your
µ , K D Y H V H H Q \ R X U E U R W K H U ¶ Morano

(a ¶) m - F F D P - W %
1 Œ € have.3 Œ € called-, ‡ • …
µ V K H K D V F D O O H G P H ¶

(b ¶)- F F D P - W a% W W X W W %
have.3 Œ € called-, ‡ • … to all
µ V K H K D V F D O O H G D O O ¶ Colobraro

) L Q D O \ W K H D O W H U Q D Q W - R F F X U V L I W K H U G S
the l- object clitic. This gives rise to subtle micro-variation depending on the contexts where l- is inserted. In the dialects we are investigating negative contexts may affect the realization of the 3rd person element. Thus, in the dialect of Morano, in negative contexts O S « occurs, (5a), while in that of Colobraro the incorporating form appears D O V R D I W H U W K H Q H J D W L R Q D V L Q D ¶

(5) (a) nu ll - 1 1 X vist-u
Neg 3 Œ € have. Œ € seen † Œ €
µ , G L G Q R W V H H K L P ¶ Morano

D ¶ L O O % Q ' Q Q a F F D P - W %
he Neg (3 ‰ Œ) have. called , ‡ • …
'he did not call K L P K H U W K H P ¶ Colobraro

7 K H G L V W U L E X W L R Q R I W K H V W H P Y R Z H O V -- and a-
L G H Q W L I \ -- as the basic allomorph, considering that it occurs in all contexts where ol
clitics are not inserted[3].

A third system is attested in the North-Calabrian dialect of Albidona, in which the enclisis on the 1st and 2nd singular persons of the verb, in (6a), coexists with the reali zation of 3rd person l- 2 & O L Q Q G D Q G U G V L Q J X O D U S H U V R Q V F 7 K H V W H P R I W K H D X [L O L D U \ K D V W K H D O W H U Q D where both alternants are admitted, i.e. the incorporation on the initial of the verbal form, or the insertion of l-. The examples in (6d, e) illustrate the 1st and the 3rd plural persons which in turn select for enclisis. (6f) illustrates the negative form, where the OCl l- R F F X U V L Q H Q F O L V L V R Q W K H Q H J D W L R Q D Q G W K by the enclitic form on the auxiliary, here -a, feminine singular.

(6) (a) - G G d X DV W %
have.1 Œ € † Œ € • Œ seen- ‰ ‡ • …
µ , K D Y H V H H Q K L P K H U W K H P ¶

(b) l - L V W %
‰ Œ have.2 Œ € seen- , ‡ • …
µ \ R X K D Y H V H H Q K H U K L P W K H P ¶

E ¶ - E E L V W %
3 ‰ Œ have. Œ seen- , ‡ • …
µ V K H K D V V H H Q K H U K L P ¶

(c) (l) D % V % L V W %
3 ‰ Œ have.2 ‰ … seen- , ‡ • …
µ < R X K D Y H V H H Q K H U K L P W K H P ¶

(d) W % m- u G D W %
Œ € have.1 ‰ … † Œ € given- , ‡ • …
µ Z H K D Y H J L Y H Q L W W R \ R X ¶

3 On the basis of the traditional measure of complexity for the rules, we could assume the substitute of 2 & O S V L L P S O H Z D \ Z K L O H a R e q u r e s a n u n J more complex list of contexts.

(e) 5 X % U % n- u Y L V W %
they have.3 ‰ … 3. † Œ € seen- , ‡ • …
µ W K H \ K D Y H V H H Q L W K L P ¶

(f) ' ll-a G G d % Y L V W %
Neg 3- • Œ € have.1 Œ € seen- , ‡ • …
µ , K D Y H Q R W V H H Q K H U ¶ Albidona

It is of note that in these dialects there is no phonological process or constraint preventing > O @ I U R P F R P E L Q L Q J Z L W K > D @ R U > - @ D V V K R Z Q E

(7) > ‰ O - U L µ Y W D K @ H J U D V V ¶
> ‰ O D W µ W Q K D @ J U D S H V ¶ Morano

> - N Z % @ µ W K H Q H H G O H ¶
> ‰ O D W µ W K Q ‰ @ U D S H V ¶ Colobraro

More to the point, l- regularly occurs before have with the lexical reading of possession, as in (8a, b).

(8) (a) l D G G d %
3 ‰ Œ have.1 Œ €
µ , K D Y H L W ¶

(b) l - M %
‰ Œ have.2 Œ €
µ < R X K D Y H L W ¶ H W F Colobraro

In other words, nothing suggests an origin of this phenomenon as due to the phono logical assimilation of l D Q G - R U a- S R W W K E H U \ W K H G L V W U L E X K \ S R W K H V L V I U R P µ E H ¶ F I 0 D Q] L a- L D U R G P 6 µ K Y D R Y L H ¶ according to syntactic properties is crucially implied. Moreover, the fact that the alter nation concerns only the auxiliary creates further obstacle to a morpho-phonological explanation. As to the lexical verbs initiating with a vowel we find different solutions. Typically, if the original initial vowel is unstressed, i.e. pre-tonic, it is deleted and the vocalic OCls are inserted, as in (9a); in some varieties the initial vowel can incorpo rate the OCl of 3rd person, as in (9b). If the initial vowel is the stressed nucleus of the stem, either l- 2 & O L V L Q V H U W H G D V L Q F R U D Q L Q V H U W H G X V X D O O \ W K H Y R L F H G Y H O D U I U L F D W L Y like in (9d) and (9e).

(9) (a) X DW W D1 N N %
† Œ € • Œ €ie.1 Œ‰ …
µ , W L H K L P K H U W K H P ¶ Albidona

(b) D V S - W W < U <
(‰ Œ) wait-3 Œ €
µ V K H Z D L W I R U K L P K H U W K H P ¶ Morano

(c) l D G G d %
‰ Œ€ave.1 Œ €
µ , K D Y H L W ¶ Albidona

(d) a 5 U - S %
• Œ €open.1 Œ €
µ , R S H Q L W ¶ Colobraro

(e) u M - S %
† Œ €open.1 Œ €
µ , R S H Q ¶ Morano

The picture we get is the following:

- have D O O R P R U as, where the latter externalizes the 3rd person internal argument in active syntax; the a- auxiliary form excludes the insertion of the object clitic.
- 3rd person OCls have the alternants: u/a/i in proclisis with lexical verbs or in enclisis on the auxiliary, l(-u/a/i L Q Q H J D W L Y H G H R Q W L F L P S H U D W L Y
- The realization of OCls is sensitive to the negative operator
- OCls have different realizations depending on the person: in Colobraro dialect the 3rd singular person and the 2nd plural incorporate the 3rd person OCl, while all other persons incorporate it; in Albidona the 2nd and 3rd singular persons select the procliticl-, contrary to the 1st singular and plural persons and 3rd plural.

7 K H D O W H U Q D W L R Q- a is related to complex of isoglosses South-Italian phenomena concerning the expression of 3rd person OCls in auxiliary contexts (see section 4; cf. Manzini and Savoia 2005, 2010). Manzini and Savoia propose a structural treatment whereby the auxiliary in C subsumes the 3rd person properties. Differently from W K D W D Q D O \ V L V Z H F R Q Q H F W W K H G L I I H U H Q W U H D O properties of such lexical items and rely on the idea that morphological complex forms are constructed on the basis of the same computational mechanism of syntax.

A crucial point is that the drop of 3rd person OCls before the auxiliary interplays with the enclitic occurrence in some varieties, like the one of Colobraro in (2a)–(2b). 6R DQ LQWHUHVWLQJ FRQMHFWXUH LV WKDW WKH FR enclisis of 3rd person OCls and, finally, their co-existence, as in (2), are the faces of the same phenomenon, implying a substantially identical morpho-syntactic procedure.

1.1 The Object Clitic System

Let us consider the clitic paradigm of these varieties. Object clitics, apart the alternants incorporated in the auxiliary seen in section 1, present three alternants:

- the simple Definiteness root l-, that we indicate with ‰ Œ, generally encompassing all referential properties, before verbs beginning in vowel as in (8);
- the complex forms where l- is combined with the inflection of gender and number, in imperative and in negative contexts (see below), indicated as 3- • Œ € † Œ € ‰ …
- the simple inflection of gender and number in proclisis, in declarative sentences, glossed as the bundle of agreement features • Œ € † Œ € ‰ …

In these dialects the plural is generally expressed by the only exponent of the 1st and 2nd person clitics, we have, as in Standard Italian and in most Italian varieties, a single form for direct and indirect object, i.e. P © 1 ‰ Œ and W © 2 ‰ Œ. The clitic forms are illus in (10a) for Morano, in (10b) for Albidona and in (10c) for Colobraro.

(10) (a) P % W % X O Y L D < U < L
1 Œ € Œ € Œ € • ‰ € .see-3 Œ €
µ V K H V H H V P H \ R X K L P K H U W K H P ¶ Morano

(b) P % W % X ‰ ‰ D è % Q %
1 Œ € Œ € Œ € • ‰ € . see-3 ‰ …
µ W K H \ V H H P H \ R X K L P K H U W K H P ¶ Albidona

(b) P % W % X ‰ Y L D % è % Q %
1 Œ € Œ € Œ € • ‰ € see-3 ‰ …
µ W K H \ V H H P H \ R X K L P K H U W K H P ¶ Colobraro

(11a), (11b) and (11c) illustrate the dative in the dialects of Morano, Albidona and Colo E U D U R D ¶ E ¶ D Q G dative/accusative clitic corresponding V W U dialects. The following distribution emerges:

- Accusative and dative of the 1st and 2nd person clitics are syncretic, presenting a single alternant in the two contexts;

- The dative in the dialect of Morano is realized by i when it is alone, as in (11b), D Q G E \ W K H V \ Q F U H W L F © Q D W L Y H d e d i W L Y H 2 & O L Q E ¶
- In Colobraro dialect the dative is realized in all contexts by the partQi©.
- The order of clitics in the string provides for the dative and 1st and 2nd person clitics before the accusative, i.e. the same as in many Romance varieties.

, Q W K H J O R V V H V 7 9 L Q G L F D W H V W K H 7 K H P D W L F 9 R Z H the root and the inflection.

(11) (a) L U ' Q D W % N N Z L V W %
‰ Œ } z • give-3 Œ € this- , ‡ • …
µ V K H J L Y H V W K L V W R K L P K H U W K H P \ R X ¶

D ¶Q u W U ' Q L U L
‰ Œ } z • 3. † Œ € give-3 Œ €
µ V K H J L Y H V L W W R K L P K H U W K H P Morano

(b) ` % è D % N N Z L V W %
‰ Œ } z • gi Œ € 3 Œ this
µ V K H J L Y H V W K L V W R K H U K L P W K H P P H ¶

E ¶ u P è D 5 %
‰ Œ } z • 3 † Œ € give.1 Œ €
µ , J L Y H L W W R K L P K H U W K H P P H ¶ Albidona

(c) Q % G - N N Z L V W %
‰ Œ. } z • J L Y H 7 9 this- , ‡ • …
µ V K H J L Y H V W K L V W R K L P K H U W K H Q ¶

(c) n u G - W %
‰ Œ. } z • 3. † Œ € J L Y H 7 9 Œ €
µ V K H J L Y H V W K L V W R K L P K H U W K Colobraro

As shown by the comparison of (10a) and (11a) for Morano, and (10b) and (11b) for Albidona, dative and accusative plural are syncretic, so that i D Q G ` % H Q F R P S D V ‰ and that of 3rd person dative. Manzini and Savoia (2017a, 2018, 2020) account for this coincidence, assuming that both plural and dative are based on the semantic relation of inclusion, C, i.e. part-whole or set. Plural implies the sub-set reading (cf. Chierchia 1997), while dative can be traced back to the elementary predicative relation where

an argument is included into or part of another, its domain (cf. Belvin and den Dikken (1997)[4] and Manzini and Savoia (2011a,b)). In other words, this analysis suggests that WKH RUGHU GDWLYH VW QG SHUVRQ ± DFFXVDWLYH the possessum, as in the case of the order possessive-DP in many languages.

A specialized alternant including the Definiteness base l- followed by the inflectional elements a, u, i µKHU KLP WKHP¶ LV LQVHUWHG ZKHQ V (cf. (5)) and in enclisis on imperatives. In both contexts, however, the order ofclitics remains the same. This allomorphy presents a stable micro-variation also attested in our sample of dialects. In the negative contexts, not all dialects introduce l- forms, as evidenced by the comparison between (12a, b) for Morano and Albidona with l-, and F IRU &RORE UDUR SUHVHUYLQJ YRFDOLF 2&OV $V for Morano, l- forms combine the definiteness root Z LW K WKH JHQGHU QXPE which in turn is however sufficient to realize the object clitics in the declarative positive contexts in (10).

(12) (a) nu OO X QµYLDU < U< OO L
Neg 3- † Œ € 33 • ‰ Œ s e e - 3 Œ €
‘(V KH GRHV QRW VHH KLP KHU WKHP¶ Morano

(b) ' OO X OO D‰ ‰ L è % Q‰ %
Neg 3- † Œ € 3- 3 Œ ‰ € … , s e e - 3 ‰ …
µWKH\ GR QRW VHH KLP KHU WKHP¶ Albidona

(c) Q ' Q QX D L ‰ Y L Q‰ %
Neg † Œ € • Œ s e e - ‰ ……
µWKH\ GR QRW VHH KLP KHU WKHP¶ Colobraro

4 The idea that all types of possession, including inalienable and psych state possession, fall under the same basic relarion resumes the analysis of possession in Belvin and den Dikken (1997, 170) according to whom “entities have various zones associated with them, such that an object or eventuality may be included in a zone associated with an entity without being physically contained in that entity… The type of zones which may be associated with an entity will vary with the entity”.

Imperatives[5] require the insertion of OCls in enclisis selecting l-V forms, as in (13a, b, c)¦l-V IRUPV RFFXU DOVR LQ FOLWLF FOXVWHUV ZRUG VWUHVV ZLWK WKH SHQXOWLPDWH YRZHO DV L

(13) (a) FD‰P -O X O D O L
FDOO 7391† Œ€ 3 3 Œ‰€…
µ&DOO KLP KHU WKHP ¶

D¶U'Q D Q‰QL l-u
JLYH 739‰Œ }z3-†Œ€
µ*LYH LW WR KHU KLP ¶ Morano

(b) FD‰PDOO %
call 3-,‡•…
µ&DOO KLP KHU ¶

E¶a- ‰P LOO %
give Œ€ 3-,‡•…
µ*LYH PH LW ¶ Albidona

(c) FD‰P DOO %
&DOO 79 ,QÀ
µ&DOO KLP KHU WKHP ¶

F¶-a- P‰P LOO %
*LYH 79 Œ€- 3-,‡•…
µ*LYH PH LW ¶ Colobraro

---

5 In Southern Italian dialects in imperatives clitic clusters attract the main stress of the word, DVVRFLDWHG ZLWK WKH ¿UVW FOLWLF KRVW e.g. UMQD Q¼Q LOX µJLYH KLP LW¶ 0RUDQR DV LOOXVWUDW
(i) F
G ?
U'QD QQL OX
In some dialects also simple object clitics following the imperative attract the word stress, as in the examples in (13a, b, c). The reassignment of the main stress in post-verbal position goes together with the insertion of 3rd person clitics with the host. Manzini and Savoia (2017a) FRQFOXGH WKDW WKH HOLFLWHG VRXQG RI WKH JUDZULWWHQ LQ elements must be read outside of the scope of the modal operator.

,PSHUDWLYHV RI VW QG SOXUDO SHUVRQ WULJJ
locative object clitics (characterizable as deictic), in contexts where the 3rd person clitic occurs in enclisis (cf. Manzini and Savoia 2017a, Baldi and Savoia 2020). In other words, mesoclisis shows up only in deictic/ dative/ locative clitics + verbal-inflection + 3rd person clitics strings, as in (14) and (15). (14a) and (15a) and (14b) and (15b) illustrate mesoclisis of deictic clitics in the contexts of 1st plural and 2nd plural forms respectively. (15c) illustrates the post-verbal position of simple object clitics. In the GLDOHFW RI $OEHGRQD VW QG SOPXUDO DQG 2QG SOXUDO DUH U
and mesoclisis[6]. As for the imperative person inflections, these dialects have a paradigm similar to the Italian: 1st and 2nd plural coincide with the present indicative forms; 2nd singular of the first class has the specialized ending -a, as in (15a) for Morano. In Albidona in final unstressed position the inflectional vowels have changed to -ə.

(14) (a) d-a- `` % ‰ P X O O %
J L Y H 7 3 9 ‰ …- 1 ‰ …- , Q À
µ / H W X V J L Y H K H U K L P W K H P L W ¶

(b) S ' U W % V B % ‰ W - O O %
E U L Q J 1 7 9 ‰ …- 2 ‰ …- ' H I , Q À
µ % U L Q J X V L W ¶ Albidona

(15) (a) U ' Q D Q ‰ Q L mu- l-u
J L Y H 7 9 3 ‰ Œ- } z 1 ‰ …- Def- † Œ €
µ / H W X V J L Y H L W W R K L P K H U ¶

(b) U ' Q D P ‰ P L ti- l-u
J L Y H 7 9 Œ €- 2 ‰ …- Def- † Œ €
µ * L Y H L W W R P H ¶ Morano

As noted above, in these systems mesoclisis is admitted only on condition that the 3rd clitic is present in final position, as in (15a, b) for Morano, so that with single OCl we KDYH HQFOLVLV DV LQ ¶D DQG ¶D IRU VW SOXU

¶(a) S ' U W D ‰ P X O O %
E U L Q J 1 7 ‰ … ' H I , Q À
µ / H W X V EU L Q J L W ¶

6 The alternant P © V © seems to be traceable back to the combination of the 1st plural exponent P © with V © WKH UHÀH[LYH LPSHUVRQDO HOHPHQW 0DQ]LQL

(b) F D P % ‰ W - P %
F D O O 7 2 9 ‰ … Œ €
µ & D O O P H ¶ Albidona

¶(a) F D – P D mu- l-u
F D O O 7 1 9 ‰ . D ef- † Œ €
µ / H W X V F D O O K L P ¶

(b) U X ‰ Q D W % P % kwiss-u
J L Y H 7 2 2 ‰ … Œ € that- † Œ €
µ * L Y H P H W K D W ¶ Morano

In negative imperatives OCls, alone or in the string dative+accusative, occur in proclitic position, between the negative marker and the verb. The 2nd singular is lexicalized by W K H L Q I L Q L W L Y H D V L Q D D Q G D D ¶ , Q W K H F D V L Q E F F ¶ D Q G E F forms in (16a) and (17a), while in Colobraro variety the simple 3rd person form is preserved, as in (18a). The strings dative+accusative D U H H [H P S O L I L H G L Q E F F ¶ D Q G D ¶

(16) (a) ' ll-u F D – P D
Neg 3- † Œ € wait- • •
µ ' R Q ¶ W F D O O K L P ¶

(b) ' `` u G D P %
Neg 3 ‰ Œ } z t- † Œ € J L Y H 7 9 ‰ …
µ / H W X V Q R W J L Y H L W W R K H U K L P W K H P ¶

(c) ' mm u G D W %
Neg Œ € † Œ € J L Y H 7 9 ‰ …
µ ' R Q ¶ W J L Y H L W W R P H ¶

F ¶ `` u G D W %
Neg 3 ‰ Œ } z † Œ € J L Y H 7 9 ‰ …
µ ' R Q ¶ W J L Y H L W W R K H U K L P W K H P ¶ Albidona

(17) (a) nu- ll-u F D – P -
Neg 3- † Œ € wait- • •
µ ' R Q ¶ W F D O O K L P ¶

D¶u- nn u U X ‰ Q -
Neg to. 3 ‰ Œ † Œ € J L Y H 7 9
µ ' R Q ¶ W J L Y H L W W R K H U K L P W K H P ¶

(b) nu- nn u U X Q - P X
Neg to.3 ‰ Œ † Œ € J L Y H 7 9 ‰ …
µ / H W X V Q R W J L Y H L W W R K H U K L P W K H P ¶

(c) nu- mm u S X U W - W L
Neg Œ € † Œ € J L Y H 7 9 ‰ …
µ ' R Q ¶ W J L Y H L W W R P H ¶ Morano

(18) (a) Q ' Q Q u F D – P - W %
Neg † Œ € call- • • ‰ …
µ ' R Q ¶ W F D O O K L P ¶

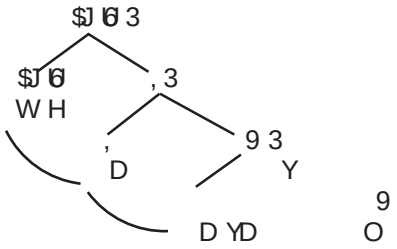
(b) Q ' mm- u G -
Neg 1 ‰ Œ lt. † Œ € J L Y H 7 9
µ ' R Q ¶ W J L Y H L W W R P H ¶ Colobraro

Summarizing:

- The enclitic form of accusatives includes the definiteness lexical base l-, missing in proclitic elements.
- Mesoclisis is triggered only by clitic clusters in 1st and 2nd forms of imperative.
- ' D W L Y H O R F D W L Y H D Q G V W S H U V R Q F O L W L-F V R F F X satives occur to the right of inflection.
- Negative imperatives require clitics and clitic clusters to be inserted between the negation and the verb, in proclisis.
- l- clitics occur when immediately preceded by the negative head.

, V 7 K H U H D 0 R U S K R O R J L F D O & R P S R Q H Q W "

In the generative syntax framework, the best known generalization concerning the distri E X W L R Q R I L Q I O H F W L R Q D O P R U S K H P H V L V % D N H U ¶ V P R Y H V W R F R P E L Q H Z L W K W K H F O R V H V W V X I I L [9 D W AgrS, that closes the complex word, as in (19), representing the 2nd plural of the Italian imperfectlava-va-te µ \ R X S O Z D V K H G ¶ 7 K H 0 L U U R U 3 U L Q F L S O syntactic operations the idea, traditional in generative grammar, that the composition of complex words is anordered cyclic mechanism. At once, it associates the treatment of inflection with syntax.



Distributed Morphology (DM) (Halle and Marantz 1993, 1994), the traditional approach to morphology within the generative framework, identifies morphology with an autonomous component, in which the insertion of morphemes is however based on an insertion mechanism in which subword elements (affixes and clitics), are understood as "dissociated morphemes" conveying information "separated from the original locus of that information in the phrase marker" (Embick and Noyer 2001, 557) and involving post-syntactic rules of linear adjacency (Local dislocation) (Embick and Noyer 2001). Hence, agreement and case morphemes are not represented in syntax but they are added postsyntactically "during Morphology". Thus, we can expect there are morphological elements devoid of any syntactic import, "ornamental pieces of PRUSKRORJ\´ DV LQ WKH FDVH RI 7KHPDWLF 9RZHOV I cf. Calabrese 2015).

As an illustration, coming back to the alternation a- YV - LQ '0 IUDPHZR FKDQJH I a-R could be seen as the outcome of a "Fusion" adjustment rule of the W\SH LQ ZKHUH WKH 3 IHDWXUHV GHILQLQJ WKH R head. As a result, the insertion of the object clitic is prevented and the specialized form of the auxiliary is inserted.

(20) [$_{OCl}$ 3 IHDWXUHV@ $X[Æ ‡ > $X[3 IHDWXUHV@

The motivation of rules such as (20) is to create the correct slot for the subsequent 9RFDEXODU\ LQVHUWLRQ 7KH TXHVWLRQ LV ZK\ OD (Hauser, Chomsky and Fitch 2002), should commit themselves to obscure the relation between sensory-motor (SM) and interpretive (IC) interfaces (Manzini and Savoia 2011a,2018). Naturally, we defend a vocabulary-based framework in the sense of Bobaljik (2002,53), that is "the pieces that constitute paradigms and rules for-gener ating them". More precisely, approaches based on abstract constructs such as basic paradigms, understood as the grammatical level organizing the morphological structure DQG DFFRXQWLQJ IRU V\QFUHWLVP :LOOLDPV D explanatory structure in the grammar (Bobaljik 2002). On the contrary, it is possible to think of the relation between syntax and morphology as based on the lexical content of those "pieces".

:H ZLOO IROORZ D GLIIHUHQW DSSURDFK WR PR
morphology is part of the syntactic computation and there is no specialized-compo nent for the morphological structure of words (Manzini and Savoia 2017b, 2011a, Manzini et al. 2020, Savoia et al. 2018). Morphosyntactic features of lexical elements, including morphemes, are fully interpretable, and contribute to externalizing the syntactic structure. Morphemes are endowed with semantic content, so excluding Late Insertion and the adjustments provided by Distributed Morphology, such as the manipulation of WHUPLQDO QRGHV LPSRYHULVKPHQW DQG IXVLRQ UX

Inflected words are analyzed as the result of a Merge operation that combines inflectional heads with a category-less lexical root R, interpreted as a predicate. In the case of nominal elements, inflectional contents as GHQGHU JHQGHU IHPLQLQH P other classificatory properties such as number and case (Manzini and Savoia 2011b). In LQIOHFWHG YHUEDO IRUPV DJUHHPHQW IHDWXUHV DQ with R. Specifically, syncretism and other kinds of ambiguity imply a treatment based RQ WKH LQWHUSUHWLYH SURSHUWLHV RI WKH LWHP V\QWDFWLF VWUXFWXUHV 6LPLODU FRQFOXVLRQV DUH DQG VSHFLILFDOO\ IRU PRUSKRORJ\ V\QWD[UHODWL

As we noticed, also subword elements are bona fide lexical entries endowed with interpretive content and contribute to forming the relevant structure. So, we can assume that the Merge operation (Chomsky 2020a,b) in (21) underlies the combination of morphemes in complex words:

(21) Merge (X,Y) Æ [X,Y]

Specifically, morphology involves the combination of heads, roots and other morphemes. Chomsky (2020a: 55) sees in pair-merge the way of treating head raising: “ , W ¶V DOZ GHVFULEHG LQFRUUHFWO\ ,I D YHUE UDLVHV WR LQI WKH 7 9 FRPSOH[EHFRPHV D 7 EXW LW¶V QRW LW¶V verbal, not inflectional.” As for modification as in the case of an adnominal adjective expression such as young man, Chomsky concludes that it is the result of an operation of conjunction where the same categorizer (Link) is shared by the conjuncts (roots) merge ZLWK WKH /LQN Chomsky refers to Heim & Kratzer (1998). Marantz (1997), speaks of categorizers such as n, that w H FDQ FRQFHSWXDOL]H DV WKH EX characterize the functional content of words entering into the agreement operations.

Drawing on Manzini (2021) and Baldi and Savoia (2021), it seems natural to assume that n is the label for the class and number features of nominal agreement. Extending this idea to verbs it is possible to identify v with the verbal categories of tense, aspect DQG PRRG WKDW PDNH DQ HYHQWLYH VWDWLYH URR inflection, for instance of tense or agreement, is sufficient to make a root, generally used as a noun, a verb, as in the case of (s)he water-s/-ed. In the model proposed here,

Agreement can be accounted for as the morphological manifestation of the identity between referential feature sets corresponding to the arguments of the sentence. In other words, there is no uninterpretable category triggering raising of a goal (see Chomsky et al. 2019, Chomsky 2020a,b).

An effective intuition proposed by Marantz (2001, 2007) is that words correspond to phases, substantially to work spaces, formed by combining the uncategorized lexical root with inner and outer morphological elements, where typically the latter are the inflections. Inflectional morphemes select for the compound including the root and its immediately attached morpheme. This model, therefore, excludes the separation between inflectional morphology, introduced in syntax, and derivational morphology, substantially OH[LFDO DV LPSOLHG LQ %DNHU¶V DQG XVXDOO\ LQ is that "syntax perform[s] all merger operations including those between morphemes within a word" (Marantz 2001, 6). A point remains to be clarified, i.e. the role of the little "v, n, a" determining "the syntactic category for roots". As we suggested above, we identify these elements with the verbal or nominal features expressed by morphemes immediately combining with the root.

3.1 Proposals for the Analysis of Clitics

Let us consider the treatment of verbal inflection and clitic insertion in terms of merger operations. Starting from Halle and Marantz (1993, 1994) the interplay between clitics DQG LQIOHFWLRQV KDV LQGXFHG D XQLILHG WUHDW associated with syntactic structure. This solution is confirmed by mesoclisis (Manzini and Savoia 2011b, Baldi and Savoia 2020), presented in (14)–(15), where clitic elements are interpolated within the inflectional string.

:H DVVXPH WKDW WKH LQIOHFWLRQDO FRQWHQW R properties of the sentence. For instance, consider the simple clause in (22) (cf. (10a), Morano):

(22) [$_{OCl}$ a] [$_{7\ Y}$ YLU <U<@
her see-3 Œ €
'(V KH VHHV KHU¶

In the light of Chomsky (2015, 2020b), the inflection, identifying the EA of the verb, is PHUJHG ZLWK 5 JLYLQJ ULVH WR D ODEHOHG DPDOJD can be identified as the realization of the category v, as in (23).

(23) < vir$_{R}$ <3UÆ[$_{v}$ $_{S}$ vir + iri $_{3}$]

If words, here the verb, are phases, we need to think that inflectional head is accessible to operations at vP, where it agrees with the features of v, as suggested in (24).

(24) (a) vP phase T v word-phase
u_3 3 L U L Y L U

More to the point, if, in accord with Roberts (2010, 57), Romance OCls have to be considered the head of agreement for v phase, the OCl can be treated as the phase edge, specifically merged with the verb realizing v, as in (25).

(25) < OCl_3, viriri$_v$ > Æ [u [viriri]] …

: H F D Q W K L Q N W K D W L Q I O H F W L R Q D O S U R S H U W L H V R
sense that vP and TP absolve the Agreement criterion invoked in Chomsky (2015, 2020b), yielding (26), where the amalgam OCl+R is associated with T.

(26) < T_3, [u [vir-iri $_3$]] > Æ [> X Y L]U < U <

In imperatives we find the order verb-OCls D V J H Q H U D O O \ L Q 5 R P D Q F H
does the order in imperative come from? In generative tradition the inversion of OCls is connected to the movement of the verb to C or to a higher position, as proposed in cartographic models where the illocutionary nature of imperatives is associated with the Speech Act Phrase (Speas and Tenny 2003), implying a directive illocutionary force as a property involved in their interpretation. In semantic literature imperatives are devoid of truth value not making assertions about the current world (Han 2011); in other words, they assign a property to a prominent argument, identified with the addressee, rather than denoting events (Platzack and Rosengren 1998, Portner 2004). In keeping with Portner (2004, 239), we can treat imperatives as predicates resulting from abstraction R S H U D W R U including an argumental variable fixed by the addressee. As suggested in (27), the imperative form U M µ Q L D Y H ¶ 0 R U D Q R L Q F O X G H V W K H
with the inflection -a of 2nd person specialized for imperative, differing from the usual ending of 2nd person that in this type of dialects is H V J b ‰ U ' Q < V < µ \ R X J L Y H

(27) U M_2 Q D µ J L Y H ¶
[J L Y H x,y), x = Addressee

& K R P V N \ L G H Q W L ¿ H V S K D V H V Z L W K O H [L F D O V X E D
SM and C-I interfaces as the result of the operation of Transfer. The procedure is constrained by the Phase Impenetrability Condition (PIC) (Chomsky (2001, 14), whereby structure [$_{ZP}$ Z…[$_{HP}$. [H YP]]], where Z and H are heads, the complement YP of H is not accessible to operations at ZP and only H and its edge are accessible to such operations (Richards 2011).

As for the enclisis of the clitic string in imperatives, we remind that it reflects a structural possibility independently implemented by the externalization in Italo-5RPDQFH YDULHWLHV DV LQ WKH FDVH RI VRPH :H Savoia 2005), that show enclisis also in declarative forms. In Standard Italian enclisis FKDUDFWHUL]HV LQILQLWLYH JHUXQGLYH DQG SDUWLF can wonder what is the link between enclisis and non-veridical contexts. The simplest hypothesis is that the left position of the verb realizes the scope of the predicative abstraction on the subject and the other arguments. In this sense, the order verb-OCls is the morphological encoding of the imperative reading.

:H FDQ WUHDW WKH 2&OV VWULQJ DV D FRPSOH[L to the IA l-u yielding the amalgam lui µWR KHU KLP WKHP LW¶ LQ D the relation of possession between the possessum, the accusative, and the possessor, the dative (Manzini and Savoia 2011b, Baldi and Savoia 2021). The cluster merges to [$_v$ UMQ D@ DV WKH UHDOL]DWLRQ RI WKH 3 IHDWXUHV IRUP RI FOLWLFV VDWLVI\ WKH UHTXLUHPHQWV RI 7 & of OCls are the result of merging of dative and accusative is supported by the evidence provided by many dialects, where dative and accusative assume specialized realizations in clusters. This is the case of Morano, where the dative is VRODWLRQ LQ E¶ it is ni in clusters, in (15a). The order dative-accusative reflects the usual order of these clitics also in preverbal position. As noticed in the previous discussion, the linearization possessor-possessum expresses the scope of the relation, where the dative restricts the UHIHUHQWLDO FRQWHQW RI WKH REMHFW FOLWLF DV genitive element.

(28) (a) < ni$_3$, l-u$_3$ > Æ [$_8$ ni [lu]]
(b) < [$_3$ ni [lu]], [$_v$ U'Q D@[$_3$ hi kE] [$_3$ U'Q D @ @ @
(c) < T$_3$, [$_v$ QL OX ÆU['Q D U@Q@ @ Q‰ QL OX @ @

Continuing along this line of analysis, if clusters are merged into specialized amalgams, we should conclude that mesoclisis is the result of a similar morphological procedure. In other words, a string as P^{-}P$_{nfl}$-lu WµWR PH \UR XQLWH P$_{nfl}$-TuQ JWYLH SO PH L from (15b) for Morano, implies that the inflection -ti$_{2pl}$ does not have a different status from that of clitics (cf. Halle and Marantz 1994), i.e. a clitic string is formed that includes the inflectional exponent, as in (29a). In (29b) the cluster is merged to the verbal stem and the verbal inflectional element -ti is externalized as usually to the left of the OCl, yielding (29b) with the effect of mesoclisis.

(29) (a) < mmi [$_3$ti], l-u > Æ [$_8$ [$_3$mmi-[$_3$ti]] l-u]
(b) < [$_3$[$_3$ P P L WL@ O$_v$> XÆ [@> UUQQ DD @ P P L WL O X @ @

At least two phenomena support the idea that OCls and the inflectional exponent are combined into a cluster, as in (29a). Firstly, the stem combining with mesoclisis is different from the one in the other contexts as highlighted by the comparison between I D W { L µ W R O W ¶ L Q W { D P G W L W X W R P H ¶ L Q E Z L W
Moreover, examples of doubling of the inflectional material are frequently realized, as in (30c) for the dialect of Albidona.

(30) (a) I D W U ti- l-u
G R 7 9 2 ‰ …- Def- † Œ €
µ G R L W W K H U H ¶

(b) I D W U Di- ti- l-u
G R 7 9 me- 2 ‰ …- Def- † Œ €
µ ' R L W W K H U H ¶ Morano

(c) d-a- t % m % ‰ - ll %
J L Y H 2 9 ‰ …-1 ‰ Œ- 1 ‰ …-it
µ * L Y H L W W R P H ¶ Albidona

The occurrence of 3rd person clitics of the type l+u satisfies a requirement of definiteness implied by these constructs, that we can relate to the non-veridicality of the imperative contexts. The idea, that we will discuss in reference to (39), is that the complex forms are inserted in order to fix referents that are interpreted independently from the scope of the modality.

The DOM effect whereby only 1st person OCls and dative clitics can occur in mesoclisis but not 3rd person clitics, can be traced back to the general point concerning the order in the clitic string. In fact, we see that pronouns interpreted in relation to the G L V F R X U V H F R Q W H [W L H V W Q G S H U V R Q F O L W L F V G L U H F W R E M H F W S U H F H G H D F F X V D W L Y H V : H K D Y H D W K H V F R S H R I W K H S D U W Z K R O H U H O D W L R Q I U R P V V 7 K X V W K H D F F X V D W L Y H G D W L Y H V \ Q F U H W L V P R I W K they are however treated as possessors independently of their thematic role (Manzini and Savoia 2010, 2011b, 2017a). This appears to be excluded for the 3rd person clitics, that however are to be read in relation to the event. In mesoclisis this order is however retained 3rd Person obliging elements in the right position. Forcing the elements of 3rd person to the right position.

3.2 The Alternation a- Y V H

As discussed in the previous paragraphs, the meaning of the sentence is projected from P R U S K R V \ Q W D F W L F S U R S H U W L H V R I O H [L F D O L W H P V

with interpretable content.[8] / H W X V F R Q V L G H U W K H D X [L O L D U \ : H
conjecture – the null hypothesis – that the auxiliary have is a verbal projection,
embedding a predicative relation between a noun and a participle selecting it as IA
(Manzini and Savoia 2011a). As for the internal structure of the participle, we see that
in Romance varieties and, specifically, in Italian dialects, inflectional properties identify
the participle with a nominal element. The category-less lexical root R combines with
the participial suffix, –t L Q W K H P R V W Y H U E D O F O D V V H V D Q G
% H W Z H H Q W K H U R R W D Q G W K H L Q I O H F W L R Q D O H O H
(31) (from (1) for Morano).

(31) [[[cam$_{R}$ -] t $_{Part}$] u $_{Infl}$] ‘called- † Œ € ¶

The participial suffix, -t- L Q K D V D U H V X O W D W L Y H V W D W L Y
7 K H P D W L F 9 R Z H O U H V X P L Q J D S U R S R V D O R I 0 D Q]
can be identified with a nominal element, introducing an indefinite variable “x”, whose
value is fixed by the internal (or external) argument of the sentence. More precisely,
thematic vowels are nominal inflections making the verbal root into a nominal form
R I W K H Y H U E D Y D L O D E O H W R L Q V H U W L Q W K H D V S H
operations, the past-participle cam- - t-u ‘called- † Œ € ¶ L V F U H D W H G E \ P H U
R in (32a), this amalgam with the stative suffix -t- L Q E D Q G W K H 3 I H D
result, in (32c). The participle in v satisfies agreement.

(32) (a) < cam$_{R}$, -$_{x}$> Æ [[$_{R}$ F D P @ - @
(b) [$_{v}$ < [$_{x}$ [cam$_{R}$ @ -$_{stative}$ @ > . W Æ [[$_{Prt}$ > W U @ P t$_{stative}$] …
(c) [$_{v}$ < [$_{Prt}$ > W U @ P t$_{stative}$, u$_{Infl}$] > Æ [$_{7}$ $_{Y}$ > > > W @ U P] u] …

Consider now in this perspective the alternant a- of the auxiliary, which is introduced
when it agrees with the participle, or, more precisely, when the participle is associated
Z L W K W K H U G S H U V R Q , $ E a \ is Label to an introduced verbal 7 K H
properties compatible with the 3rd IA specified by the participle, in (32); it works like
W K H F R U U H V S R Q G L Q J 2 & O L Q R W K H U D W W X [W K H Y H F
as an internally inflected stem selected in combination with the past participle of which
it specifies the IA. The head D W W X is pair-merged with the participle yielding (33a)
E D V H G R Q V K D U L Q J F R P S D W L E O H L H Q R Q F R Q W U

8 This assumption excludes the hypothesis that a morpheme, i.e. a non-readable category, alternates with l- in auxiliary contexts where the 3rd person is realized through a specialized form of the auxiliary.

9 For the sake of clarity, we remind that in this dialect the ending R I W K H ¿ U V W S H U V
D X [L O L D U \ L V W K H X V X D O Y H U E D O L Q À H F W L R Q F R U U H V S R Q

same argument. a 1 1- is merged to T forming the verbal amalgam a 1 1-u, agreeing with the subject, in (33b).

(33) (a) < a_3 1 1 F D> PÆW> DX $_R$ Q1> F D P] - W X
(b) [$_T$ D 1 1> … XÆ > D $_{Q1}$ 1] 1 X

The result is the sequence in (34), where the participle externalizes by the alternant a- W K H 3 S U R S H U W L H V R I Y

(34) [[$a_{3\,\langle\}}$ 1 1 X] [[9 3 [$_{Prt}$ F D P - W X Morano

The other option, implemented by the dialects of Colobraro (cf. (2b)) and Albidona (cf. (6a, c, d)), is that 3rd person clitics are in enclisis. In this respect, we recall that V W Q G S H U V R Q 2 & O V R F F X U L Q S U R F O L V L V R Q W K H D

(35) (a) P % Q % Y L V W %
1 Œ € have.3 ‰ seen- , ‡ • …
µ W K H \ K D Y H V H H Q P H ¶

(b) W % G G d X DG D W % L
Œ € have.1 Œ €† Œ € • Œ € given- , ‡ ‰ ….
µ , K D Y H J L Y H Q L W W K H P W R \ R X ¶ Albidona

E ¶ n - Q X GD - W % L
Œ € have.1 Œ €† Œ € • Œ € given- , ‡ • .‰ …
µ W K H \ K D Y H J L Y H Q L W W K H P W R \ R X ¶

(b”) n - Q X GD - W % L
3 ‰ Œ } z • have.1 Œ €† Œ € • Œ € given- , ‡ • .‰ …
µ W K H \ K D Y H J L Y H Q L W W K H P W R K Colobraro P W K H P

7 K H S U R F O L V L V R I W K H V W Q G S H U V R Q 2 & O V L V Z K D tion of OCls in declarative sentences, seen in (24). The issue is the occurrence of the 3rd person OCl in enclisis. Superficially, the distribution is similar to what we saw for imperatives in (28), where the OCl of 1st person precedes the inflectional morpheme of the verb and the 3rd person OCl in final position.

Let us assume that 3rd person OCls are pair-merged in the workspace of v, where they realize the IA agreement of v, as in (36a). In other words, the endings -u/-a/-i are to be considered as the objective inflections of the auxiliary in v, in (36a), agreeing with the participle. Then the person clitic is merged to G G Š X µ , K D Y H L W ¶ I U R P

the second object (the recipient), in (36b); this amalgam is merged to T where realizes the subject agreement, in (36c). The sequence in (36d) is the result:

(36) (a) [$_{v}$ G G, d u $_{3>}$] Æ [G G d] X
Aux 3 Aux
(b) [$_{T}$ W % > G . G d E X > W % > G G d X @
3 Aux Aux
(c) [$_{T}$ - G G u d T $_{1SG}$ > … Æ > - G - G d]
1SG 3 1SG 3 , Q I D 7
G > > W % G] G d v X @ . G D W % @ @ @
, Q I O - 9 3 L 9 3 ‰ ‹ •

The morphological link between the enclitics elements and the auxiliary is evidenced by an interesting phenomenon, i.e. the occurrence of -u/-eu- in the negative contexts where the proclitic realization of the OCl is triggered. Thus, as shown in (37), the OCl is divided into two elements, the definiteness root L(-) S U R F O L V L V D Q G W K H J element in enclisis.

(37) ' O O % G G d X Y X L V W % L
Neg 3 ‰ Œ have. Œ € † Œ € see Œ € , ‡ • … ‰ …
µ , K D Y H Q R W V H H Q K H U K L P W K H P ¶ Albidona

Two Phase contexts are realized, i.e. v and T; in the latter OCls are merged to the verb (auxiliary), as suggested in (38):

& 3 Y 3 S K D V H T V Y word-phase
Aspectual
O % K D Y H X 3 Y L V W %
3

Our intuition is that the auxiliary in the aspectual construct realizes the event properties of v, and the enclitic element behaves exactly like inflection on the IA of 3rd person, W \ S L F D O O \ L Q V H U W L Q J L W V H O I L Q I L Q D O S R V L W L R Q are anchored to the event, to the effect that they are interpreted with respect to it. On W K H F R Q W U D U \ V W Q G S H U V R Q F O L W L F S U R Q R X Q V X Q L Y H U V H V R W K D W W K H \ D U H I U H H W R R F F X S \ W K H the auxiliary realizes the aspectual properties of v, to account for its ability to introduce the reference to the IA. In other words, differently from lexical verbs, auxiliary forms have as their only content the features associated with v, both the aspectual properties and, possibly, the IA agreement features. As a consequence, we see that the auxiliary can, someway, realize the reference to IA by its inflection. The different interpretive P H F K D Q L V P R I U G S H U V R Q D Q G V W Q G S H U V R Q H [S agreement, as far as their content is not read in relation to the event but is anchored to the universe of discourse, determining a strong effect of DOM. In other words, they are introduced by independent specialized exponents.

This explanation seems to be supported by deontic periphrasis have-to-V, where the auxiliary selects for the infinitive of a lexical verb preceded by the preposition µ W R ¶
In some dialects, i.e. Colobraro and Albidona, Z H I L Q G W K H - V W H P V S U H
3rd person OCl l- L Q D D ¶ 2 W K H U V \ V W H P V G R Q R W G L I I
introducing however the stem a, as in Morano, in (39b).

(39) (a) l - dda F D O P D
3 ‰ Œ have.3 Œ € 3 ‹ ~ ‰ F D O O 7 9
µ V K H K D V W R F D O O K L P K H U W Colobraro ¶

D ¶ - dda F D O P D
3 ‰ Œ have.3 Œ € 3 ‹ ~ ‰ F D O O 7 9
µ V K H K D V W R F D O O K L P K H U W K H Albidona ¶

(b) a-r a l -
have-3 ‰ Œ 3 ‹ ~ ‰ G R 7 9
µ V K H K D V W R G R L W ¶ Morano

Interestingly, the other auxiliary constructs, as for instance the pluperfect, present a certain degree of variability, in the sense that in some dialects the auxiliary includes the reference W R W K H U G S H U V R Q D V L Q D is inserted ¶, as in Z K (40c) H L Q R V

(40) (a) D Y L M % F D P - W %
(3 ‰ Œ) have- Œ ‰ • F D O O 7 9 3 • ‰ - , ‡ • …
µ , K D G F D O O H G K H U K L P W K H P ¶

D ¶ D Y H U % F D P - W %
(3 ‰ Œ) have- & ^ ‡ } 1 Œ € O O 7 9 3 • ‰ - , ‡ • …
µ , Z R X O G K D Y H F D O O H G K H U K L P W Colobraro ¶

(b) avi-a F D P - W X
(3 ‰ Œ) av- Œ ‰ • F D O O 7 9 3 • ‰ - † Œ €
µ , K D G F D O O H G K H U K L P W K H P ¶ Morano

(c) l D Y L % Y L V W %
3 ‰ Œ have. , ‰ • s Œ € 3 • ‰ - , ‡ • …
µ , K D G V H H Q K H U K L P ¶ Albidona

: H F D Q H [S H F W W K D W D G L D O H F W V X F K D V W K H
enclisis and l-proclisis, adopts the latter solution in contexts where enclisis is

non morpho-phonologically admitted, as in the case of the imperfect. This solution appears also in the 2nd plural person of the present perfect where again the dialects vary from inserting l- (Albidona), cf. O D Y L V © F D P D W © µ \ R X K D Y H F D inserting l- (Colobraro, Morano). After all, the definiteness root for the 3rd person OCl is the outcome that occurs regularly before lexical verbs beginning in vowel, including have of possession, that excludes the alternation and behaves like the other lexical verbs, as in (9c).

$ V U H J D U G V W K H L Q D E L O L W \ R I W K H I R U P - W R U object, we point out that in these varieties only one auxiliary form is attested, that applies to all verb classes (transitives, unergatives, unaccusatives), save to represent with the stem alternant a- the 3rd person IA (% D O G L D Q G 6 D Y R L D , Q R W K H U the typical Elsewhere distribution, only registering the usual T agreement and excluding the v argumental properties (perhaps reminiscent of the properties of

: H V W L O O K D Y H W R R W K H U [illegible] Q H J D W L Y H [illegible] 3 S O U person OCls manifest a referentially richer alternants including the definiteness root l- and the J H Q G H U Q X P E H U L Q I O H F W L R Q D V L O O X V W U D W H G L Q Manzini and Savoia (2017b) propose that the introduction of the definiteness root l- (and possibly the stress, cf. fn. 1) in non-veridical contexts, such as imperative and negation (Giannakidou 1998, 2011), is required as it provides a complete referential content in contexts where the pronoun is out of the scope of the relevant operator. Combining with a non-veridical state of affairs, 3rd person OCls incorporate also the definiteness morpheme. In other words, this make them able to be interpreted independently of the X V X D O F R Q Q H F W L R Q Z L W K W K H H Y H Q W R Q D S D U Z L tive contexts the pronoun provides the restriction for the variable introduced by the negation, something like (41) (Baldi and Savoia 2021).

(41) > C x[$_{Neg}$ nu [x [ll $_{3}$]]] [$_{T}$ - 1 1 X « Morano
µ , K D Y H Q R W « L W K H U K D L P W K H P ¶

& R Q F O X G L Q J 2 E V H U Y D W L R Q V

The data concerning the realization of the 3rd person OCls in auxiliary contexts presented in Manzini and Savoia (2005, § 5.11) can help us to highlight some generalizations. If only by focusing on the data from Southern Italian dialects, we note that the realization of the auxiliary have with the simple object clitic drop is very widespread. In that corpus, we find the alternation between ‡ and the 2nd singular in Montesano (South Campania) dialect in (42a) and between h- and l- L Q V J D Q G S O X U D O L Q W K H G L D Apulia) in (42b). The systems we have found in Lausberg area are attested in other adja F H Q W Y D U L H W L H V & R O R E U D U R ¶ V V \ V W H P F K D U D F W H and Cersosimo, and the alternation between D Q G - I R U P V D V L Q W K H G L D characterizes other North Calabrian varieties, for instance that of Nocara.

(42) (a) D G G d % F D P D W %
O H F D P D W %
D F F D P D W % «
µ , K D Y H F D O O H G K L P H W F ¶ Montesano

(b) O H M % F D P D W %
D F D P D W %
D F D P D W %
O H P % F D P D W %
O H W % F D P D W %
O H Q Q % F D P D W %
µ , K D Y H F D O O H G K L P K H U W K H P H W9R¶W X L Q R

The table in (43) schematizes the distribution of the different alternants, where ‡ indicates the simple drop of the 3rd person OCl, a- the specialized form of the auxiliary, l- the insertion of the prevocalic form of the clitic, enclisis the enclitic occurrence of the OCl.

(43)	Many dialects	Morano	Montesano	9R O W XC L Q R	Colloquio	Albidona
1sg	‡	a	‡	l-	encl	encl
2sg	‡	a	l-	‡	encl	l-
3sg	‡	a	‡	‡	a	l-
1pl	‡	a	‡	l-	encl	encl
2pl	‡	‡	‡	l-	‡	l-
3pl	‡	a	‡	l-	encl	encl

:H VHH WKDW WKH VLPSOH GURS LV WKH EDVLF VROXWL particular, enclisis on the second person implies its occurrence on the first, that seems to be the specialized context for it to be implemented. Generally, the third singular person excludes HQFOLVLV VW VLQJXODU SHUVRQ DQG VW UG- SHUVRQ tion of the enclitic or l- clitic. The 3rd singular generally excludes the independent realization RI WKH SURQRXQ the realization of the 3rd person OCl can be uniform (a-) along the paradigm or not. In this second case, the overt realization of the 3rd person IA is generally associated with the persons that are discourse-implicated, such as 1st singular, or discourse anchored, 1st, 2nd and 3rd plural. Needless to say, we are speculating on constraints that, on a par with the DOM effects, belong to more external linguistic knowledge of the VSHDNHU DQG QRW WR WKH LQQHU JUDPPDU FI %REDO of the ordinary syntactic combinatory procedure of pair-merge and the effect of “third factor” constraints that regulate the interpretation (Chomsky 2005). A plausible hypothesis is that the interpretive work can possibly involve general semantic constraints.

Summing up, we have applied a morphological model based on the idea that there is no specialized morphological component nor different nature for morphological rules. The hypothesis we have pursued is that morphemes (lexical and functional) are endowed with interpretable properties that determine they occurrence in syntax, thus detaching ourselves from the typical perspective of DM. This approach has in the operation of (pair-)merge its basic mechanism, able to treat the formation of complex words and their relation with syntax.

:RUNV &LWHG

Baker, Mark. 1988, Incorporation: A Theory of Grammatical Function Changing. Chicago: University of Chicago Press.

Baldi, Benedetta, and Leonardo M. Savoia. 2019. “Auxiliary Selection in Italian Dialects: Person Split, OCls and Raddoppiamento Fonosintattico.” Quaderni di Linguistica e Studi Orientali QULSO 4: 55–105.

Baldi, Benedetta, and Leonardo M. Savoia. 2020. “Clitics in Imperative: Proclisis, (QFOLVLV DQG 0HVRFOLVLV LQ $OEDQLDQ DQG LQ Area.” LINGBAW 6: 17–46.

Baldi, Benedetta, and Leonardo M. Savoia. 2021. “Partitives and Indefinites: Phenomena LQ ,WDOLDQ 'Studia Linguistica KWWSV RQOLQHOLEUD IXOO VWXO

Belvin, Robert, and Marcel Den Dikken. 1997. “There, happens, to, be, have.” Lingua 101: 151–183.

%REDOMLN -RQDWKDQ ' 6\QFUHWLVP ZLWKRX 1981, 1994.” In Yearbook of Morphology 2001, edited by Geert Booij and Jaap van Marle, 53–85. Dordrecht: Springer.

&DODEUHVH $QGUHD ,UUHJXODU 0RUSKRORJ\ DQ Isogloss: 69–102.

&KLHUFKLD *HQQDUR 3DUWLWLYHV 5HIHUHQ In Proceedings of Semantics and Linguistic Theory 9, , HGLWHG E\ $UP 73–98. Ithaca: CLC Publications.

Chomsky, Noam. 1972. Studies on Semantics in Generative Grammar. The Hague: Mouton.

Chomsky, Noam. 2001. “Derivation by Phase.” In Ken Hale: A Life in Language, edited E\ 0LFKDHO . H–54. Cambridge, Mas.: The MIT Press.

Chomsky, Noam. 2005. “Three Factors in Language Design.” Linguistic Inquiry 36: 1–22.

Chomsky, Noam. 2015. “Problems of Projections.” In Structures, Strategies and Beyond. Studies in Honour of Adriana Belletti, edited by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 3–16. Amsterdam: John Benjamins.

Chomsky, Noam. 2020a. The UCLA Lectures (April 29 – May 2, 2019). KWWSV OLQJ DXI QHW OLQJEX]]

Chomsky, Noam 2020b. “Puzzles about Phases.” In Linguistic Variation: Structure and Interpretation: Contributions in Honor of M. Rita Manzini, edited by Ludovico Franco and Paolo Lorusso, 163–167. Berlin: De Gruyter.

Chomsky, Noam, Ángel Gallego, and Dennis Ott. 2019. “Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges.” In Generative Syntax: Questions, Crossroads, and Challenges, Catalan Journal of Linguistics Special Issue, edited by Ángel Gallego and Dennis Ott, 229–261.

& R O O L Q V & K U L V D Q G Towards a Theory of Morphology as Syntax, online. Accessed 22 June, 2022. K W W S R U G L Q D U \ Z R U N L Q J J U D P P D U L towards-theory-of-morphology-as-syntax.html (13.07.2021).

Embick, David. 2010. Localism versus Globalism in Morphology and Phonology. Cambridge, MA: The MIT Press.

Embick, David, and Rolf Noyer, 2001. “Movement Operations after Syntax.” Linguistic Inquiry 32 (4): 555–595.

Giannakidou, Anastasia. 1998. Polarity Sensitivity as (Non)Veridical Dependency. Amsterdam: John Benjamins.

Giannakidou, Anastasia, 2011. “Negative and Positive Polarity Items.” In Semantics: An International Handbook of Natural Language Meaning Y R O H G L W H G von Heusinger, Claudia Maienborn, and Paul Portner, 1660–1712. Berlin: Mouton de Gruyter.

Halle, Morris, and Alec Marantz. 1993. “Distributed Morphology and the Pieces of Inflection.” In The View from Building 20 H G L W H G E \ . H Q Q H W K + D O - D \ . H \ V H U ± & D P E U L G J H 0 D V V 7 K H 0 , 7 3 U H V + D O O H 0 R U U L V D Q G $ O H F 0 D U D Q W] ³ 6 R P H . H \ ) H D In Papers on Phonology and Morphology, edited by Andrew Carnie, Heidi Harley, and Tony Bures, 275–288. MIT Working Papers in Linguistics 21

Han, Chung-hye. 2011, “Imperatives.” In Semantics Y R O H G L W H G E \ . O D X V Claudia Maienborn, and Paul Portyner, 1785–1804. Berlin: De Gruyter Mouton.

+ D X V H U 0 D U F ' 1 R D P & K R P V N \ D Q G : 7 H F X P V H K) / D Q J X D J H : K D W , V , W : K R + D V , Volume 298: 569–1579.

Lausberg, Heinrich. 1939. Die Mundarten Südlukaniens. Niemeyer: Halle.

Manzini, M. R L W D ³ & K R P V N \ ¶ V / L Q N V D Q G / L Q N H U di Linguistica e Studi Orientali. QULSO online 7: 89–102.

Manzini, M. Rita, and Leonardo M. Savoia. 2005. I dialetti italiani. Morfosintassi-generativa Y Y $ O H V V D Q G U L D G H O O ¶ 2 U V R

Manzini, M. Rita, and Leonardo M. Savoia. 2010. “Syncretism and Suppletion in Clitic Systems: Underspecification, Silent Clitics or Neither? Syntactic Variation. The Dialects of Italy H G L W H G E \ 5 R E H U W D ' ¶ $ O H V V D Q G U R Roberts, 86–101. Cambridge: Cambridge University Press.

Manzini, M. Rita, and Leonardo M. Savoia. 2011a. Grammatical Categories: Variation in Romance Languages. Cambridge: Cambridge University Press.

Manzini, M. Rita, and Leonardo M. Savoia. 2011b. “Mesoclisis in the Imperative: Phonology, Morphology or Syntax?” Lingua 121: 1101–1120.

0DQ]LQL 0 5LWD DQG /HRQDUGR 0 6DYRLD D ³Romance: Allomorphies and (Re)Ordering.” Transactions of The Philological Society 115: 98–136.
Manzini, M. Rita, and Leonardo M. Savoia. 2017b. “Gender, Number and Inflectional &ODVV LQ 5RPDQFH) HLQJHU 8VH DQG /LQJXLVWLF 6WUXFWXUH, Q edited by Joseph Emonds and Markéta Janebová, 263–282. Olomouc: Palacký University Olomouc.
Manzini, M. Rita, and Leonardo M. Savoia. 2018. The Morphosyntax of Albanian and Aromanian Varieties. Berlin: De Gruyter Mouton.
Manzini, M. Rita, Leonardo M. Savoia, and Benedetta Baldi. 2020. “Microvariation and Macrocategories: Differential Plural Marking and Phase Theory.” L’Italia Dialettale 81: 189–212.
0DUDQW] $OHF ³1R (VFDSH IURP 6\QWD['RQ¶W the Privacy of Your Own Lexicon.” University of Pennsylvania Working Papers in Linguistics 4: 201–25.
Marantz, Alec. 2001. Words : & &) / ; ; + DQGRXW 86&) HEUXDU\
Marantz, Alec. 2007. “ 3 K D V H V D Q G 3 K D V H V ,Q W K H 7 K H R U \ R I *U D P P D U ,Q Q H G ited by Sook-Hee Choe, 191–222. Seoul: Dong-In Publishing.
Platzack, Christer, and Inger Rosengren. 1998. “On the Subject of Imperatives: A Minimalist Account of the Imperative Clause.” The Journal of Comparative Germanic Linguistics 1: 177–224,
Portner, Paul. 2004. “The Semantics of Imperatives within a Theory of Clause Types.” In SALT ; , 9 H G L W H G E \ 5 R E H U W < R X Q J ± , W K D F
Richards, M. D. 2011. “ ' H U L Y L Q J W K H (G J H : K D W ¶ V 1 H [W 4 X D 3 K D V H
Roberts, Ian. 2010. Agreement and Head Movement. Cambridge, Mass.: The MIT Press.
Savoia, Leonardo M., and M. Rita Manzini. 2010. “Lexicalization of 3rd Person Object Clitics: Clitic Enclisis and Clitic Drop.” In Syntactic Variation. The Dialects of Italy H G L W H G E \ 5 R E H U W D ' ¶ $ O H V V D Q G U R $ G D P / H G J Cambridge: Cambridge University Press.
Savoia, Leonardo M., Benedetta Baldi, and M. Rita Manzini. 2018. “Sigmatic Plurals LQ 5RPDQFH 9DULHWLHV 6SRNHQ LQ ,WDO\ DQG LINGBAW 4: 141–160.
6SHDV 3HJJ\ DQG &DURO 7HQQ\ ³&RQILJXUDWLRQ In Asymmetry in Grammar, edited by Anna DiSciullo, 315–344. Amsterdam: John Benjamins.
:LOOLDPV (GZLQ ³5HPDUNV Linguistic Inquiry 8: 101–139.
:RRG -LP DQG $OHF 0DUDQW] ³7KH ,QWHUSUH In The Verbal Domain H G L W H G E \ 5 R E H U W D ' ¶ $ O H V V D Q G U R J. Gallego, 255–278. Oxford: Oxford University Press.

Part II. Explorations in Syntax

Passivization of Multiple Complement Verbs in English

Tamás Csontos

Károli Gáspár University of the Reformed Church in Hungary

csontos.tamas@kre.hu

Abstract: In the present article I investigate multiple complement verbs in passive constructions in English. I account for the fact that recipients in the dative construction and themes in the double object constructions do not undergo passivization, as sentences like *To John was given water, *John was given water to or *Water was given John are ungrammatical. I am going to approach the issue from a new perspective. The framework that I adopt is Syntax First Alignment (Newson 2010), which assumes no constituent structure and operates with a limited set of constraints.

Keywords: passivization, multiple complement verbs, Syntax First Alignment, semantic decomposition

1. Introduction

It is a generally accepted fact that both the direct and the indirect object can undergo passivization in English. This is illustrated by (1a) and (1b):

(1) (a) Water was given to John.
(b) John was given water.

The standard assumption is that (1a) is related to the dative construction, while (1b) is related to the double object construction, compare (2a) and (2b), respectively.

(2) (a) Someone gave water to John.
(b) Someone gave John water.

From a standard point of view, the two constructions are distinct although they share certain common underlying properties. The following questions arise: why is it impossible to passivize the recipient in the dative construction (with or without the preposition) and why is it impossible to passivize the theme in the double object construction? In other words, why are the sentences in (3) ungrammatical?

(3) (a) *To John was given water.
(b) *John was given water to.
(c) *Water was given John.

The main goal of the present paper is to answer these questions. I assume that the two constructions in (1) come from almost identical sources, the only difference between them is that the direct object in (1b) is a focussed element. First of all, I am going to discuss the relevant background assumptions, following Grimshaw (2005) and Ramchand (2008). Secondly, I introduce Syntax First Alignment (SFA), i.e. the framework within which I address the issue of multiple complement verbs. Lastly, I will explain certain word order phenomena and present the analysis itself in detail.

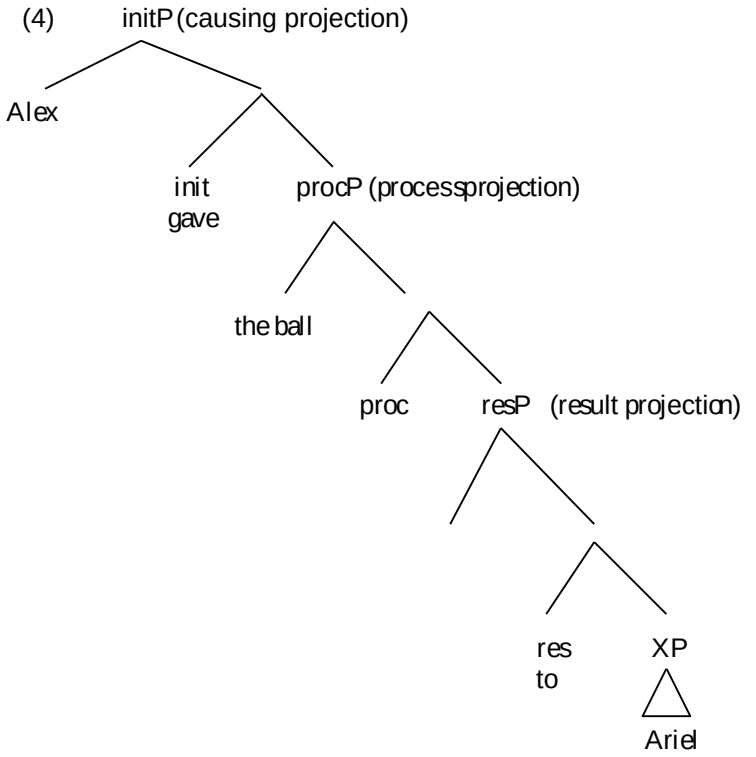
2. Background Assumptions

2.1 Semantic Decomposition and Event Structure

Following Grimshaw (2005) and using Levin and Rappaport Hovav's (2005) terminology, I assume that verb meanings are composed of root content (idiosyncratic information) and their event structures, i.e. what type of event they denote. In other words, it is proposed WKDW WKHUH LV D URRW HOHPHQW ¥ DV ZHOO DV R and argument structure related to this root.

Ramchand (2008) – among others – discusses the semantic decomposition of predicates which may contain three subevents: a causing subevent, a process-denoting subevent and a result subevent. She introduces three projections: the causing projection (headed by init), the process projection (headed by proc) and the result projection (headed by res). According to Ramchand, init is very similar to v, which licenses the external argument. Accordingly, proc and res have similar functions, they license the internal arguments. The former specifies the nature of the change or process and licenses the subject of the process (i.e. the undergoer/theme), while the latter licenses the subject of the result, e.g. the recipient.

When analysing double object verbs, as in Alex gave the ball to Sam, she proposes the following structure. Note that to is inserted under res.



In my analysis, I will refer to init, proc and res as v_1, v_2 and v_3,[1] respectively and as licensors collectively.

2.2. The Rudiments of Syntax First Alignment

The model Syntax First Alignment (SFA), which is based on Alignment Syntax, was introduced by Newson (2010). It is an approach that rejects the existence of constituent structure. In this respect, it is similar to other models such as Word Grammar (Hudson 1984) and Dependency Grammar (Debusmann 2006), but differs greatly from these non-constituent structural approaches, as Syntax First Alignment is based on general Optimality Theoretic assumptions (Prince and Smolensky 1993).

Also, it is proposed that the input consists of sub-lexical elements, called Conceptual Units (CUs), i.e. universal stock of basic units –this is an assumption which is shared by Nanosyntax (Starke 2009) as well. According to Newson and Szécsényi (2012),

1 In this respect, I follow Newson (2014).

two types of CUs can be distinguished: roots that carry descriptive semantic content[2] U H S U H V H Q W H G Z L W K W K H V \ P E R O ¥ D Q G I X Q F W L R Q content, e.g. tense, aspect, person and number. There are functional CUs, referred to as markers, which relate arguments to the event structure given a predicate. In Syntax First Alignment, the input is the basis of semantic interpretation in the same way as it is the basis of the syntactic interpretation. The input itself contains information that is relevant for both aspects of processing.

The generator (GEN) imposes linear orderings on the input elements while it is not allowed to add any element which is not part of the input. This also means that the candidate set is finite[3] – as opposed to Optimality Theory. On the other hand, there may be elements that are present in the input but are absent from the output violating faithfulness constraints – see below.

These orderings constitute the candidate set which will be evaluated by alignment and faithfulness constraints – to be discussed shortly. Only after the optimal candidate has been determined does lexical insertion occur. In other words, vocabulary items are inserted post-syntactically. In SFA, the vocabulary contains phonological forms, the associated ordered sets of conceptual units as well as the context of insertion. For example, the vocabulary entry for the progressive is as follows:

(5) L Q J < > S U[4] R J @ ¥

The notion of late lexical insertion is an integral component of more recent models as well, e.g. Distributed Morphology (Halle and Marantz 1993) or Nanosyntax (Starke 2009). According to Newson and Szécsényi (2012), four principle restrict late lexical insertion.

First, in line with Nanosyntax, they opt for the so-called Superset Principle[5] which requires that a vocabulary item can be inserted if it is associated with all the CUs which can be found in a given sequence of CUs although it may contain extra CUs as well. For instance, if the sequence to be realized is <x,y,z>, and the possible vocabulary items

2 The roots constitute what are traditionally referred to as verbs, nouns, or adjectives depending on the context which they are positioned in. For instance, a root aligned to a determiner will be realised as a noun.

3 This makes computing and evaluation much easier, which is another advantage of the present model.

4 The symbol to the left of the arrow can be pronounced as “associated with”. It represents the S K R Q R O R J L F D O I R U P R I W K H S U R J U H V V L Y H V X v [+ R Z H Y H U the orthographical forms in my paper.

5 Caha (2009) also argues in favour of the Superset Principle instead of the Subset Principle, which Distributed morphology operates with.

that can spell it out are associated with <x,y>, <x,y,z,w> and<x,y,w>, the best fitting match will be <x,y,z,w> even though it is associated with an extra <w> conceptual unit: it contains all the features of the sequence <x,y,z>.

In addition, Newson and Szécsényi (2012) further assume that only contiguous sequences can be realized by a single vocabulary item. Also, they propose that vocabulary insertion is root centric, which means that the process begins with the roots, spelling them out with those contiguous functional units which the vocabulary entry allows for, while the remaining conceptual units will be realized separately. Lastly, the principle of Minimal Vocabulary Access requires that a string of conceptual units be spelled out by the fewest number of vocabulary items possible.

The architecture of Syntax First Alignment is presented below:

LQSXW: *(1: FDQGLGDWH VHW: (9$/: RSWLPDO FDQ
;
semantic interpretation

The notion of a domain needs to be introduced, as it also plays a key role in SFA. Newson (2010, 32) defines domains as "sets of input elements which share a given property". For instance, the argument domain consists of arguments related to the same predicate. Importantly, domains are not structural units, and are not necessarily contiguous strings either, as their members can be separated by other elements. The concept of a domain is useful if the position of a given element needs to be determined with regard to two or more elements.

Let us return to the constraints mentioned above. There are only two families of constraints[6] operating in Syntax First Alignment: faithfulness and alignment constraints. Faithfulness constraints are violated if an element which is present in the input is missing from the output. In other words, these constraints guarantee that the input and the output are identical. They are generally ranked high; otherwise, many elements may be deleted from the output, which would prevent the hearer from recovering the intended meaning.

As far as alignment constraints are concerned, three basic relationships can be distinguished: precedence, subsequence and adjacency. These constraints are responsible for the position of target elements with regard to hosts, which can be single elements or domains.

(7)	(a)	xPy 'x precedes y'	violated by y...x order
	(b)	xFy 'x follows y'	violated by x...y order
	(c)	xAy 'x is adjacent to y'	violated by every CU which intercedes between x and y

6 Actually, these constraints constitute a more limited set of constraints than the ones proposed in Optimality Theory, where nothing restricts what can be a possible constraint.

When the host is a domain, the constraints can be defined in terms of the violation conditions below, following Newson and Szécsényi (2012):

(8) (a) xPDy ‘violated by every member of domain y which precedes x’
(b) xFDy ‘violated by every member of domain y which follows x’
(c) xADy[7] ‘violated by every member of domain y which is not adjacent to x’

In addition, there are also anti-alignment constraints with respect to a given domain. For instance, x*PD$_y$ requires that x cannot precede domain y. This anti-alignment constraint is violated if x precedes all the members of domain y or if D$_y$ has no members at all.

In the next sections, I am going to demonstrate that it is possible to explain the phenomena introduced in section 1 within the framework of Syntax First Alignment using only alignment and faithfulness constraints.

3.1 Basic Word Order in English

To account for the word order in passive sentences, it is necessary to briefly introduce the constraints which determine the position of the arguments, licensors, the root and the inflections.

Let us first start with the argument domain (D$_A$), which is composed of the argument makers related to a given verbal root – following Newson (2013). In English the order R I W K H V H P D$_1$ @N H >@V ! J$_2$ >@ U D K H$_3$ @ U D K H$_1$ @H $_2$ >@D UDJQ$_3$ G@ >DUUHJ D V V R F L D with what is traditionally referred to as the external argument and the two internal arguments, respectively. Newson (2013) proposes the following constraints, which are responsible for their order:

> D$_1$ @J3$_A$ ' ! > D$_2$ @UJ3$_A$ ' ! > D$_3$ @UJ3$_A$ '

The first constraint, for instance, is violated by every member of D$_A$ which precedes > D$_1$ @J

In English, the verbal root follows the subject while preceding all the other arguments. In other words, it occupies the second position in the argument domain. This second position phenomenon can be captured by the combination of an anti-precedence and a precedence constraint: the former guarantees that the verbal root does not precede the argument domain while the latter requires the verbal root to precede the argument domain. The result is that the root is not first, but as close to first as it can be (i.e. second).

---

7 The constraints in (7c), (8a), (8b) and (8c) are gradient constraints, which means that they F D Q E H Y L R O D W H G W R G L u H U H Q W G H J U H H V 7 K H R W K H U V violated gradually: they are either violated or not.

¥ 3 '! ¥ 3 '

The licensor domain (D) comprises v, v and v.[8] The order of these elements is determined by the constraints in (11):

>v@ 3 '! >@ 3 '! >@ 3 '

Before determining the position of the licensors with respect to the verbal root, it is necessary to take a look at multiple complement verbs, as they are associated with three arguments and consequently with three licensors as well.

(12) Someone gave water to John.

If we assume that it is v that is responsible for the presence of to in (12), v is not necessarily adjacent to the root. However, I agree with him that v and v are adjacent to the root and are normally spelled out with it. It seems that it is more important for v to be adjacent to > D @ than for > D @ to be as close as possible, i.e. directly following v, cf. the hierarchy of the relevant constraints (13c) and (14d) – in the table below. Note that in English the first argument precedes its licensor, while the second and the third argument follow their licensors. This can be achieved by the constraints below:

(13) (a) v A ¥
(b) v A ¥
(c) v A ¥

(14) (a) > D @ J P
(b) > D @ J F
(c) > D @ J F
(d) > D @ J v

Table (1) shows the hierarchy and the interaction of these constraints:

[8] The list of the members of this domain will be extended.

:	(15a)					**		
	(15b)					*		
	(15c)		*(!)			***		
	(15d)					*** (!)	*	
	(15e)				*(!)	**		*

Table 1. Order of arguments, licensors and the root

As demonstrated above, the winning candidate is (15a), as it best satisfies the given constraints. Candidate (15b) loses on the highest-ranked constraint, as the third argument is not adjacent to v_3, which is a fatal violation – marked with '!'. (15c) and (15e) are also ruled out, because they violate higher-ranked constraints: the former violates the constraint which requires the third argument to follow v_3, whereas the latter violates the constraint which guarantees the adjacency of v_2 and the root. Candidate (15d) loses on the constraint which requires v_3 to be adjacent to the root: this constraint is violated three times, as three CUs intercede between them, while (15a) violates this constraint only twice.

I assume that the vocabulary entry for give is as in (16), while to is associated with v_3 in its vocabulary entry (c.f. 17):

J L Y H $_1$ ¥ $_2$ Y v_3

W R $_3$ < Y

Note that the Superset Principle allows the sequence v_1-v_2 in (15a) to be spelled out by give as well. In this case, v_3 needs to be realized independently by to, because only contiguous sequences can be spelled out by a single vocabulary item.

Now let us take a look at (18), where the recipient directly follows the verb.

(18) Someone gave John water.

The only difference between (12) and (18) is the presence of the focus CU associated with the direct object in (18) which is responsible for the position of this object, as focussed objects tend to follow the argument domain.[9] This is ensured by a focus constraint:

(19) Foc F D_A

This constraint is ranked lower than $>_{IO}$ and $>_{Adv}$, as focussed subjects typically fronted, but is ranked higher than $>_{D_2}$ and $>_{D_3}$:

$>_{D_1}$ ≫ $>_{D_2}$ ≫ Foc F D_A ≫ $>_{D_3}$

Recall that the root takes the second position in the argument domain. The constraints which are responsible for this are in (10). Tableau (2) illustrates how these constraints yield the desired word order.[10]

		¥ 3 '	¥ 3 '	> D_1	Foc F D_A	> D_2	> D_3
(21a)	> D_1 > I R F_2 D > U J_3		*		*(!)	*	**
(21b)	> D_1 > D_3 U J > I R_2 D U J		*			**	*
(21c)	¥ > D_1 J > I R_2 D *(!) U J					**	*
(21d)	> D_3 > D_1 J > I R_2 D U J		*	*(!)		**	
(21e)	> I R_2 F D > J_3 D > D_1 U J		*	**(!)	**		*

Table 2. Order of arguments in double object constructions

9 This is in line with Rochemont and Culicover (2009), who state that structural foci in English are right-peripheral. (Note that this also accounts for the fact that sentences like *I gave the beautiful girl it are ungrammatical. Pronouns are not associated with focus, as they typically represent old information. Therefore, the IO>DO order where the indirect object DP and the direct object is a pronoun is not motivated.) Secondly, a similar observation is made by Røreng (2011), who investigates the order of arguments in German. She claims that the canonical word order in German is DO>IO (just like in (QJOLVK ZKLFK FDQ EH PRGL¿HG E\ WKH IRFXV IDFWRU IRFX

10 The licensors are not included in this table, as my focus is now on the order of the arguments and the root. I will return to this issue in the next section.

11 For the sake of convenience, I use the CU [foc]$_2$ (focussed second argument) as D VKRUWKDQG@IRU > IRF@ > DUJ

Note that the proposed constraints can also account for the order of arguments and the root in (1b), repeated as (22). Bear in mind that it was assumed that the direct object is a focussed element while the first argument is not present, as it lacks an agent.

(22) John was given water.

:	(23a)	> D 3 @ J > H R F 2 @ D U J		*			*	
	(23b)	> D 3 @ J > I R F 2 @ D H J		**(!)			*	
	(23c)	H > 3 @ D U J R F 2 @ D U J		*(!)			*	
	(23d)	H > I R F 2 @ D U J 3 @ J		*(!)		*		*
	(23e)	> I R F 2 @ H U J 3 @ U J		*		*(!)		*
	(23f)	> I R F 2 @ D U 3 @ U H J		**(!)		*		*

Table 3. Order of arguments in passive constructions with a focussed direct object

The last domain which plays an important role in establishing the word order in English is the inflection domain (I), which is made up of the inflectional, the perfect, the progressive and the passive conceptual unit. As the order of these CUs is fixed, the following constraints are introduced (Newson 2013):

> , @ 3 ' > S H U H @ S ' U R J @ S ' D V V @ 3 '

As the examples in (25) demonstrate, the root must precede the last element of the inflectional domain. This can be guaranteed by the constraints in (26).

(25) (a) He has been singing
> , @ > S H U H @ S U R J @

(b) It had been being built
> , @ > S H U H @ U R H @ S D V V @

H) ' ! H) '

Another important issue is the position of the inflections with respect to the licensors. I assume that v_1 and v_2 are adjacent to the root under normal circumstances and thus spelled out with it.

However, v_3 is less connected to the root, which (12), repeated as (27) below, demonstrates.
7 K H L Q I O H F W L R Q D O H O H P H Q W $, realized by to: > S D V V @ L V F O R V H U

(27) Someone gave water to John.

The conclusion is that the constraint which requires inflections to be adjacent to the
U R R W L H > L Q I O @ $ ¥ $ L ¥ V D U Q $ ¥ [illegible]

(28) v A ¥ ! A ¥ ! > L Q I O @ $ ¥ ! Y

3.2 Passivization and Multiple Complement Verbs

It is often argued that passive sentences contain an implicit agent. In Csontos (2017), I propose that the passive morpheme -*en* spells out two CUs: the implicit general external argument (GEA) and the passive CU (> S D V V @ Z K L F K O L F H Q F H V L a licensor, i.e. a member of D. Note, however, that > S D V V @ L V D P H P E H U R I W domain as well. The vocabulary for the passive -*en* is as follows.

H Q < > S D V V @ > * ($ @

The question is why the preposition is absent in (18) and (22) but present in (27) and in (1a), repeated as (30):

(30) Water was given to John.

Importantly, the CU which the preposition realizes, i.e. v_3, must be present in the inputs: there is no semantic difference between the dative construction (where v_3 is spelled out by *to*) and the double object construction (where it is seemingly absent) - apart from the focus on the direct object in the latter. Similarly, there is no semantic difference between (22) and (30) either, besides the fact that the direct object is focussed in the former. So, the question arises as to what happens to v_3 in (18) and (22). My claim is that it gets deleted from the output - violating faithfulness constraint. More specifically, I identify the presence of the focus on the second argument as the condition on which v_3 is deleted[12]. This can be achieved by the following constraints:

(31) v) > I R F @ D U J > I R F @ D U J (v [13]

12 , I W K H G L U H F W R E M H F W L V Q R W I R F X V V H G W K H V H F R Q of the relevant CUs will be determined by lower ranked constraints.

13 % R W K R I W K H W Z R K L J K H U U D Q N H G F R Q V W U D L Q W V D U violated by the reverse order of the given elements.

Note that deletion is generally problematic from an interpretability point of view. However, v_3 is an element that does not need an independent realisation for the result to be interpretable. Therefore, licensors are more easily deleted than, for example, arguments themselves, as they are recoverable from the visible argument and the content of the root.

Now let us see how the assumptions and the constraints introduced so far can yield the desired results. Let us start with (22), *John was given water.

Candidates 32a. > D3 @ J > S D V V @ > * ($ @ L V U H D O L] H G @ : D W H U
32b. > D3 @ J > S D V V @ > * ($ @ L V U H D O L] H G @ : D W H U
32c. > D3 @ J > S D V V @ > * ($ @ L V U H D O L] H G @ : D W H U
32d. > D3 @ J > S D V V @ > * ($ @ L V U H D O L] H G @ : D W H U
32e. v3 > D3 @ J > S D V V @ > * ($ @ L V U H D O L] H G @ : D W H U

	¥)'I	¥)'I	v3) > IR2F@	v3 3 > IR2F@	Faith (v3)	> D3@J3v3$	> D3@J3) Y	v2A ¥	> LQA¥@	v3A ¥	Foc F DA	> D2@J3'A	> D3@J3'A
(32a)		*	*(!)			***	*		**	*		*	
(32b)		*			*				*			*	
(32c)		*		*(!)		6*	*		*	****		*	
(32d)		*	*(!)			5*	*		*	***		*	
(32e)		*	*(!)										

Table 4. Order of CUs in passive constructions with multiple complement verbs involving a focussed direct object

The optimal candidate is (32a), where the root and v_2 are spelled out by the recipient John and the theme water, respectively, while the root and v_3 are realized by the verb give ± IROORZLQJ WKH 6XSHUVHW 3URFHGXUH, whereas the > SDVV@ > * ($ @ LV UHDOL]HG DV V\QWKHWLF SUSKHP tion of why *John was given water (c.f. 32c) and *To John was given water (32e) are ungrammatical: (32b) is a more optimal candidate.

As tableau (5) demonstrates, v_3 is realized separately if the theme is not focussed, as in Water was given to John:

Candidates a. > D 2 @ J > S D V 2 V 3 > @ S D ¥ V @ > > D 3 @ S @

b. > D 2 @ J > S D V 2 V 3 @ S @ D V ¥ @ V @ > > * D 3 @ S @ @ Y

c. > D 2 @ J > S D V 2 V 3 @ S @ D V ¥ @ V @ > 3 * @ S @ @ > D U J

d. > D 2 @ J > S D V 2 V 3 @ S @ D V ¥ @ V @ > 3 * @ S @ @ > D U J

	¥) ' I	¥) ' I	v 3) > I R 2 F @ D U	v 3 3 > I R 2 F @ D U	Faith (v 3)	> D 3 @ V 3 $	> D 3 @ J 3) Y	v 2 A ¥	> L Q A ¥ @	v 3 A ¥	Foc F D A	> D 2 @ J 3 ' A	> D 3 @ J 3 ' A
(33a)		*				* * (!)	*		**	*			*
(33b)		*							*	***			*
(33c)		*			*(!)				*				*
(33d)							*(!)		*	****			*

Table 5. Order of CUs in passive constructions with multiple complement verbs

The winning candidate is (33b). As the table illustrates, v_3 needs to be realized separately by to, as the verb cannot spell it out with the root and v_2, although it is associated with all the licensors in its vocabulary entry, see (16). The reason for this is that these elements do not form a contiguous sequence, as v_2 and v_3 D U H V H S D U D W H G E \ > > * ($ @ 1 R W H D O V R W K D W Z H F D Q D F F R X Q W I R U W K H John (c.f. (33c): (33b) is more optimal.

4. Conclusion

In the present paper, I have demonstrated that it is possible to account for the presence and absence of the preposition to in passive constructions involving multiple complement verbs in English by using verbal/semantic decomposition, alignment and faithfulness constraints and late lexical insertion. It has been argued that it is the focussed direct object which is responsible for the missing preposition both in active and passive sentences: if the direct object is focussed, the CU which spells out, i.e. v_3 is deleted from the

14 The examples in table (4) and (5) may suggest that v_3 is never realized by give, as this licensor either gets deleted if the theme is focussed or it is spelled out independently. This would also mean that it is unnecessary to include this CU in the vocabulary entry for this verb – as opposed to what is claimed in (16). Note, however, that there are examples, where the verb must realize v_3, e.g. I gave him it, as it cannot be deleted: recall that pronouns are not usually focussed elements. Consequently, give must be associated with v_3 in its vocabulary entry.

output – violating a faithfulness constraint. In other cases, is_3 is present in the output and realized separately.

Works Cited

Caha, Pavel. 2009. “The Nanosyntax of Case.” PhD diss., University of Tromsø.

Csontos, Tamás. 2017. “Word Order and Passivization in English and German: A Syntax First Alignment Approach.” PhD diss., Eötvös Loránd University.

Debusmann, Ralph. 2006. “Extensible Dependency Grammar: A Modular Grammar Formalism Based on Multigraph Description.” PhD diss., University of Saarland.

Grimshaw, Jane B. 2005. Words and Structure. Stanford: CSLI Publications.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger H G L W H G E \ . H Q Q H W K / + D O H D Q G 6 D P Cambridge, MA: The MIT Press.

Hudson, R. 1984. Word Grammar. Oxford: Blackwell.

Levin, Beth, and Malka Rappaport Hovav. 2005. Argument Realization. Cambridge: Cambridge University Press.

Newson, Mark. 2010. “Syntax First, Words After: A Possible Consequence of Doing Alignment Syntax without a Lexicon.” In The Even Yearbook 9, edited by László 9 D U J D í % X G D S H V W (| W Y | V / R U i Q G 8 Q L Y H U V L W http://seas3.elte.hu/delg/publications/even/2010.html#nm

Newson, Mark. 2013. “What You Can Do with Domain and Why They Are Not Phrases.” In VLlxx, edited by Péter Szigetvári, n.p. Budapest: Eötvös Loránd University. Accessed September 14, 2021. http://seas3.elte.hu/tmp/vlfs/newson.html.

Newson, Mark. 2014. “Binding Revisited.” In The Even Yearbook 11, edited by László 9 D U J D í % X G D S H V W (| W Y | V / R U i Q G 8 Q L Y H U V L W http://seas3.elte.hu/delg/publications/even/2014.html

Newson, Mark, and Krisztina Szécsényi. “2012. Dummy Auxiliaries and Late Vocabulary Insertion. In The Even Yearbook H G L W H G E \ / i V] O y 9 D U J D Eötvös Loránd University. Accessed September 14, 2021. http://seas3.elte.hu/delg/publications/even/ 2012.html#nm.

Prince, A., and P. Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. New Jersey: Rutgers University Center for Cognitive Science.

Ramchand, Gillian Catriona. 2008. Verb Meaning and the Lexicon: A First-Phase Syntax. Cambridge: Cambridge University Press.

Rochemont, Michael Shaun, and William Peter Culicover. 2009. English Focus Constructions and the Theory of Grammar. Cambridge: Cambridge University Press.

Røreng, Anita. 2011. “Die Deutsche Doppelobjektkonstruktion.” PhD diss., University of Tromsø. Accessed September 14, 2021. https://munin.uit.no/bitstream/handle/10037/3377/thesis.pdf?sequence=2&isAllowed=y.

Starke, Michal. 2009. “Nanosyntax – A Short Primer to a New Approach to Language.” In Nordlyd 36.1, edited by Peter Svenonius, Gillian Ramchand, Michal Starke DQG .XWW 7DUDOGV HQ í 7URPV¡ &$67/ http://www.ub.uit.no/baser/nordlyd/.

Gradients of Reflexivity: Psych Verbs in Causative Alternations

3 U H G U D J . R Y D Ï H Y L É

University of Novi Sad, Serbia

predrag.kovacevic@ff.uns.ac.rs

Abstract: This paper addresses the issue of the syntactic and semantic representation of anticausatives with SE. The “reflexive analysis” extends the reflexive semantics of SE to anticausative constructions (Chierchia 2004), while the “standard view” is that anticausatives are versions of transitive constructions without the external argument (Schäfer and Vivanco 2016). Based on the tests from event modifier licensing (Alexiadou et al. 2014 inter alia), I show that a significant portion of psych verb anticausatives pattern with typical reflexives and transitives in licensing instrumental NPs/DPs introducing the causer argument whereas typical anticausatives license *od* (‘from’)-PPs. I interpret this data as evidence for an intermediary category between reflexives and anticausatives that I label “semi-reflexives”. Contrary to the “reflexive account”, the analysis presented in the paper (drawing on Sportiche 2014) does not ascribe the same structures and denotations to reflexives and anticausatives, while in contrast to the “standard view”, it still maintains a link between these different uses of SE.

Keywords: psych verbs, anticausatives, reflexives, causative alternations, Serbian

1. Introduction

The issue of the correct syntactic and semantic representation of anticausative constructions has received a significant amount of attention in the recent literature. The examples in (1) from Serbian (1a–1a’) and English (1b–1b’) illustrate two broader types of anticausatives – those involving morphological marking (the morpheme SE in Serbian) and those without any morphological changes of the verbal stem (the English causative alternation). One line of research maintains that anticausatives are akin to reflexives pointing to the fact that many languages employ the same morpheme (SE) to derive reflexives

and anticausatives (Chierchia 2004; Koontz-Garboden 2009, 2012). The opposing view holds that anticausatives are merely structurally impoverished versions of transitive verbs lacking the external argument (Parsons 1990; Schäfer and Vivanco 2016).

Psych verbs (verbs denoting emotional states or events – e.g. amuse or love) are a class of verbs that exhibit numerous syntactic and semantic peculiarities (Belletti and Rizzi 1988; Pesetsky 1996; Landau 2009). Some of these puzzling behaviors are observed in the domain of reflexives and in the formation of anticausatives. Belletti and Rizzi 1988 note the "reversed binding patterns" with psych verb reflexives. Levin $ O H [L D G R X D Q G , R U G F K L R D L D D Q G R W K H U J H Q H U D O \ G R Q R W S D U W L F L S D W H L Q F D X V D W L Y H D O languages such as Greek, Romanian or Serbian, where causative alternations are asso- F L D W H G Z L W K R Y H U P R U S K R O R J L F D O P D U N L Q J R Q W K examples). In light of these facts, exploring psych verb anticausatives, in languages in which they are available, could yield important insights into the syntax and semantics of anticausatives more generally.

(1) (a) Toma je istopio ãHüHU causative transitive
Toma. ‡ ˆ † z Ž ‘ melt. ‰ z Œ sugar. ‡ ˆ †
'Toma melted the sugar.'

(a') âHüHU se istopio anticausative
sugar. ‡ ˆ † Œ ~ melt. ‰ z Œ •
'The sugar melted.'

(b) Tom caramelized the sugar. causative transitive

(b') The sugar caramelized. anticausative

(2) (a) Toma je zadivio Anu causative transitive
Toma. ‡ ˆ † z Ž ‘ amaze ‰ z Œ Ana. ‡ ˆ †
'Toma amazed Ana.'

(a') Ana se zadivila anticausative
Ana. ‡ ˆ † Œ ~amaze ‰ z Œ •
'Ana got amazed.'

(b) Tom amazed Ana. causative transitive

(b') *Ana amazed. anticausative

The aim of this paper is to analyze the internal syntactic structure of psych-verb anticausatives (involving the morpheme SE) in Serbian to determine what (if anything) makes them different from "ordinary/typical" anticausatives. If psych verb SE anticasuatives are syntactically and/or semantically different from "ordinary/typical" anticausatives, then, this difference might tell us something about the lack of these structures in languages such as English. The structural differences between psych verb anticausatives and typical anticausatives will be established by testing for their combinability with various types of event modifiers (Alexiadou and Anagnostopoulou 2009; Alexiadou et al. 2014; Alexiadou et al. 2015; Gehrke 2013, 2015; Gehrke and Marco 2014). It will be shown that psych verb anticausatives can license instrumental case-marked NPs/DPs introducing the causer (3a) while ordinary anticausatives license od('from')-PPs instead (3b).

(3) (a) Ana se oduševila Tominom pesmom.
Ana. ‡ ˆ † z Ž ' amaze‰ z Œ T oma. ‰ ˆ Œ Œ p e Œ ; ‡ Œ •
'Ana was amazed by Toma's poem.'

(a') â H ü H U se karamelizovao od toplote.
Sugar. ‡ ˆ † Œ ~ caramelize‰ z Œ from heat. € ~ ‡
'The sugar caramelized from heat.'

In this respect, psych-verb anticausatives pattern with reflexives, even though they lack the agentivity component which is necessarily present with reflexives. To capture these facts, I will propose a graded or ranked model of reflexivity based on the options supplied by the extended VP structure. Specifically, I will argue that (i) pure reflexives arise when SE combines with full-fledged Voice; (ii) "semi-reflexives" arise when SE combines with v^0 with a filled Spec position; and (iii) typical anticausatives arise when SE combines with v^0 without the Spec position. The discrepancy between those psych-verb anticausatives that license instrumental NPs/DPs and typical anticausatives will be attributed to the structural distinction in the higher layers of the extended VP structure. Namely, it will be argued that the presence of an external argument either in Spec vP or Spec Voice P is responsible for licensing instrumental NPs/DPs with reflexives and "semi-reflexive" psych verb anticausatives (as well as agentive transitives). On the other hand, the absence of the external argument with typical anticausatives (and unaccusatives) will be taken as the reason behind the incompatibility between these structures and instrumental NPs/DPs requiring the use of od('from')-PPs to introduce the cause participant.

The paper is structured as follows. Section 2 introduces the opposing views regarding the correct structural representations of anticausative structures and points to the potential significance of anticausatives derived from psych verbs in this debate. In Section 3, I introduce the diagnostics of the presence of particular layers of the extended

VP structure based on the combinability with different types of even modifiers. These diagnostics will be used to show that psych-verb anticausatives are structural different from typical anticausatives. The structural representations modelling the observed differences are introduced in Section 4. Section 5 concludes the paper.

2. Anticausative SE: Reflexivizer or Pure Anticauzativizer

The literature on the status of SE can (roughly) be divided into two camps. So-called "reflexive accounts" always treat this morpheme as a reflexivity marker (Chierchia 2004; Koontz-Garboden 2009, 2012). A more mainstream line of research, the so-called "standard account" (Schäfer and Vivanco 2016) denies any kind of synchronic syntactic or semantic link between the reflexive SE and the anticausative SE (Parsons 1990; Schäfer and Vivanco 2016). The first two parts of this section will present the arguments of the two approaches to the semantic and syntactic contribution of SE. The third part of this section will briefly illustrate the cross-linguistic variation when it comes to the availability of anticausatives derived from psych-verbs.

2.1. Reflexive Accounts

As is transparent from the label attached to this set of accounts, "reflexive accounts" extend the reflexive semantics of SE to constructions typically labelled anticausative (4b). Chierchia (2004) proposes to interpret (4b) along the lines of "the house toppled/collapsed itself" (i.e. as a reflexive). More precisely, following Chierchia's (2004) analysis, one would interpret the sentence in (4b) as entailing that some property of the house was such that it led to its collapse.

(4) (a) Ana je srušila N X ü X
Ana. ‡ ˆ † z Ž ‘ topple.‰ z Œ • house. z | |
'Ana toppled the house.'

(b) . X ü D se srušila
house. ‡ ˆ † Œ ~ topple.‰ z Œ •
'The house collapsed'

The precise semantics that Chierchia (2004) proposes for anticausatives presupposes that the introduction of the morpheme SE builds on the basic semantics of the transitive verb by identifying the external argument with the internal argument as with typical reflexive verbs.

One piece of evidence Chierchia (2004) offers in support of his analysis concerns the distribution of da se ('by itself') with anticausative forms. Namely, while this expression is licensed in agentive transitive constructions (5), it is blocked with non-agentive transitives (6), passives and impersonals (6), but it is allowed with anticausatives (7). The

fact that SE anticausatives pattern with agentive transitives is expected and accounted for assuming that the syntax and semantics of the anticausative version contains the semantics of the transitive version.

(5) Gianni mi ha picchiato da se agentive
Gianni me has hit by self
'Gianni hit me by himself.'

(6) (a) *Gianni conosce il latino da se stative verb
Gianni knows the Latin by self of cognition
'Gianni knows Latin by himself'

(b) *Gianni ha sudato da se verb of physical
Gianni has sweat by self function
'Gianni sweat by himself.'

(c) A: Tu hai fatto sudare Gianni contextually improved
You have made sweat Gianni verb of physical function
'You made Gianni sweat.'

B: No, ha sudato da se
No has sweat by self
'No, he sweat by himself.'

(7) (a) *La porta e stata aperta da se passive
the door is been opened by self
'The door was opened by itself.'

(b) *Questo libro si legge da se impersonal
this book one. | .. reads by self
'This book reads by itself.'

(8) (a) La porta si e aperta da se anticausative
the door Œ ~is opened by self
'The door opened by itself.'

(b) La barca e D u R Qda D W D se
the boat is sank by self
µ 7 K H E R D W V D Q N E \ L W V H O I ¶

One desirable conceptual consequence of this analysis is pointed out by Koontz-Garboden (2009, 2012). According to Koontz-Garboden (2009, 2012), assuming that anticausatives are a sub-type of reflexives avoids violating the Monotonicity Hypothesis (Kiparsky 1982), according to which additional morphological material can only add, but not remove meaning. If the addition of this morpheme simply builds on the basic semantics of the transitive verb, it follows that morphological complexity correlates with semantic complexity as predicted by the Monotonicity Hypothesis. Alternatively, if the role of SE was simply to remove the external argument, the addition of a morpheme would result in semantic (and syntactic) impoverishment contrary to the predictions of the Monotonicity Hypothesis.

Koontz-Garboden (2009, 2012) also points out that the anticausative construction is not entailed by its transitive causative counterpart, which is what one would expect if the role of the anticausative construction involved only a subset of the denotation of the transitive one. The lack of the entailment relationship between the causative transitive form and the anticausative one is illustrated with the Spanish example in (9). This example is significant because the first portion of the sentence contains the anticausative form of one verb under negation while the second part affirms the causative transitive form of the same verb.

(9) El vaso no se rompio, lo rompiste tu
the vase not Œ ~ broke it broke you
'The vase didn't break, you broke it.'

According to Koontz-Garboden (2009, 2012), the fact that one can deny the anticausative construction in the first part of the sentence in (9) while affirming the transitive one in the second part without creating a fatal contradiction shows that the anticausative construction is not entailed by the transitive one.

2.2. Standard Account

In contrast to the "reflexive account" of SE, what is sometimes referred to as the "standard account" argues precisely that the anticausative construction is nothing but the version of the transitive construction without the external argument. Schäfer and Vivanco (2016) claim that the semantic representation of anticausative verbs lacks the [CAUSE] component as well as the external argument in contrast to the transitive counterpart. Therefore, the semantics of the transitive construction properly contains the semantics of the anticausative one, the only difference being that the external argument is left unexpressed in the anticausative version.

In response to Koontz-Garboden's (2009, 2012) treatment of sentences like (9), Schäfer and Vivanco (2016) argue that these examples represent instances of linguistic negation rather than logical negation. In essence, they argue that what is

negated in the first part of the example in (9) is not the entailment of the clause but the conversational implicature enabled by the anticausative construction. According to these authors, by negating the first part of the sentence in (9) the speaker actually denies that it is a sufficiently informative description of the situation at hand (i.e. they are accusing the interlocutor of violating the Gricean Maxim of Quantity).

Countering the "reflexive analysis" of anticausatives, Schäfer and Vivanco (2016) point out that in some anticausative constructions, any kind of reflexive semantics would create non-sensical denotations (10).

(10) A gap opened

The argument here is that applying the reflexive analysis of SE to the anticausative construction in (10) would result in a denotation whereby the gap caused itself to come to existence, but, of course, no entity can be the cause of its own existence. In contrast, this issue does not emerge with the "standard account" because the denotation of the sentence in (10) would include just an unexpressed cause rather than identifying the cause with the theme.

2.3. Psych Verbs in Causative Alternations

Psych verbs are particularly interesting in relation to the question of anticausative and reflexive constructions because they exhibit numerous syntactic and semantic peculiarities (Belletti and Rizzi 1988; Pesetsky 1996; Landau 2009). Famously, (the vast majority of) English psych verbs generally fail to derive anticausative constructions (Levin 1993). The verb annoy in (11a) is a causative transitive psych verb with the experiencer argument in the object position (a so-called "object experiencer verb") denoting a change of mental state on the part of the experiencer. However, unlike other causative transitive (change-of-state) verbs, psych verbs cannot be used as anticausatives as shown in (11b).

(11) (a) The movie annoyed John
(b) *John annoyed

In contrast, languages that allow SE or other anticausativity markers exhibit no such restrictions, as exemplified by Romanian (12a) and Greek (12b). The example (12a) from Romanian shows a pair of object experiencer verbs (i.e. causative transitives), which are used intransitively (i.e. as anticausatives) in combination with the reflexive morpheme SE. The Greek example in (12b) shows one causative transitive psych verb, which is also used intransitively but without the morpheme SE. Instead, so-called "non-active" morphological marker is added to the verbal stem.

(12) (a) Ion s-a V X S U D W v Q de U L L a V W e a W cea Q R X în
John ‹ •-has upset/worried of at news the new in
cinci minute/ repede / în parc asta a avut loc ieri
¿ Y H minutes quickly in part this has had place y.day
µ - R K Q J R W X S V H W Z R U U L H G I U R P W K H Q H Z V L Q W K

(b) O Janis stenahorithike me ta nea se pende lepta/
the John saddened. ‡ z | •with the news in ¿ Y H minutes
grigora sto parko afto sinevi htes
quickly in park this happened yesterday
µ - R K Q T X L F N O \ J R W V D G I U R P W K H Q H Z V L Q ¿ Y H
7 K L V K D S S H Q H G \ H V W H U G D \ ¶

Since anticausative psych verbs are possible in other languages such as Greek, Romanian or Serbian, the question is why English psych verbs do not participate in causative alternations given the existence of equivalent constructions in other languages.

What is also worth noticing in (12) is that Romanian and Greek use formally different expressions to introduce the causer participant. In Romanian (12a), it is a *de*('of/from')-PP while in Greek, the causer is introduced by means of a *me*('with')-PP typical of instruments (12b). This will be significant in Section 3 where it will be shown that the equivalents of these two types of PPs are both possible with SE anticausatives in Serbian, but with different types of verbs.

2.4. Summary

The "reflexive" and "standard" accounts of anticausatives predict different syntactic and semantic behaviors of these constructions. The "standard view" predicts unaccusative properties for anticausatives while the "reflexive view", of course, predicts reflexive properties for anticausatives, presumably that means that they have an external argument. Next, since English psych verbs do not participate in causative alternations, there is a reason to believe that in languages where psych-verb anticausatives exist, they might exhibit behaviors that would help us tease apart the predictions of the two competing approaches to the syntax and semantics of anticausatives, which will be the aim of the remainder of this paper. Specifically, the existing tests for the presence/absence of different layers of verbal structure will be applied to psych-verb anticausatives in Serbian in order to determine whether or not they show signs of the higher portions of verbal structure (vP, VoiceP) predicted to be missing under the 'standard account' but present under the "reflexive account".

3. Diagnosing the Internal Structure of Serbian Psych Verb Anticausatives

One way of diagnosing the presence of the different layers of verbal structures in various derivations comes from the licensing of various types of event modifiers. By-phrases are assumed to be licensed by VoiceP (Kratzer 1994; Gehrke 2013, 2015; Alexiadou et al. 2014; Alexiadou et al. 2015) explaining why passives (13a) but not anticausatives (13b) can license these items.

(13) (a) The car was broken by Mary
(b) *The car broke (down) by Mary

Typical anticausatives (14a) and unaccusatives (14b), on the other hand, license from-PPs expressing the causer.

(14) (a) The car broke (down) from the cold temperature
(b) The patient died from a heartattack

PPs expressing instruments tend to be licensed in active agentive constructions (15a) and with passives (15b) but not with anticausatives (15c).

(15) (a) Mary opened the door with a key
(b) The door was opened with a key
(c) *The door opened with a key

The Serbian equivalent of the English by-phrase is the strane (‘from side of’)-PP and it is licensed with passives but not with anticausatives (16).

(16) (a) Vrata su otvorena od strane provalnika
door. ‡ ˆ † z Ž ‘ open. ‰ z Œ Œfro‰ ‹ • side. € ~ ‡ burglar. € ~ ‡
‘The door was opened by the burglar.’

(b) *Vrata su se otvorila od strane provalnika
door. ‡ ˆ † z Ž ‘ Œ ~open. ‰ z Œfrom side. € ~ ‡ burglar. € ~ ‡
Literally: ‘The door opened by the burglar.’

Instrumental case-marked NPs/DPs are licensed with transitive active constructions (17a) and passives (17b) but not anticausatives (17c).

(17) (a) Marija je otvorila vrata N O M X þ H P
Marija. ‡ ˆ † z Ž ‘ open. ‰ z Œ • door key. , ‡ Œ •
‘Marija opened the door with a key’

(b) Vrata su otvorena N O M X þ H P
door. ‡ ˆ † z Ž ‘ open. ‰ z Œ Œ key ‰ ‡ Œ •
‘The door was opened with a key’

(c) *Vrata su se otvorila N O M X þ H P
door z Ž ‘ Œ ~ open. ‰ z Œ • key. , ‡ Œ •
Literally: ‘The door opened with a key’

Finally, od(‘from’)-PPs introducing the causer are licensed with anticausatives (18a) but not with active transitives (18b) and passives (18c).

(18) (a) Vrata su se otvorila od vetra
door. ‡ ˆ † z Ž ‘ Œ ~ open. ‰ z Œ • from wind. € ~ ‡
‘The door opened from the wind.’

(b) *Marija je otvorila vrata od vetra
Marija. ‡ ˆ † z Ž ‘ open. ‰ z Œ • door. z | | from wind. € ~ ‡
Literally: ‘Marija opened the door from the wind.’

(c) *Vrata su otvorena od vetra
door. ‡ ˆ † z Ž ‘ open. ‰ z Œ Œ from ‰ ‹ wind. € ~ ‡
Literally: ‘The door was opened from the wind.’

Turning now to the domain of psych-verb anticausatives, we can apply these tests to determine whether these constructions pattern with reflexives or with typical anticausatives and unaccusatives. Here, we encounter a more complex picture. By-phrases are always rejected (19).

(19) (a) *Marija se ohrabrila od strane brata
Marija. ‡ ˆ † Œ ~ encourage ‰ z Œ • from side brother. € ~ ‡
Literally: ‘Marija got encouraged by her brother.’

(b) *Marija se iznervirala od strane brata
Marija. ‡ ˆ † Œ ~ annoy. ‰ z Œ • from side brother. € ~ ‡
Literally: ‘Marija got annoyed by her brother.’

(c) *Marija se zaprepastila od strane brata
Marija. ‡ ˆ † Œ ~ shock. ‰ z Œ • from side brother. € ~ ‡
Literally: ‘Marija got schocked by her brother.’

(d) *Marija se razbesnela od strane brata
Marija. ‡ ˆ † Œ ~ anger. ‰ z Œ • from side brother. € ~ ‡
Literally: ‘Marija got angry by her brother.’

If the rejection of by-phrases were taken as a decisive diagnostic for the lack of VoiceP layer, then, one would have to assume that psych verb anticausatives, like “typical anti-causatives” always lack this portion of the extended VP structure. However, one should be cautions not to jump to this conclusion prematurely since there are agentive structures, which presumably include the VoiceP domain, but systematically reject by-phrases. For instance, typical reflexives disallow by-phrases; however, they do show signs of the presence of the external argument on other tests such as control into purpose clauses (Gehrke 2013, 2014; Alexiadou at al. 2014).

Indeed, with some psych verbs, SE triggers a purely reflexive reading as evidenced by the availability of control into purpose clauses (20a) and the grammaticality of a reflexive pronoun *sebe* (‘self’) (20b).

(20) (a) Marija se hrabri [da bi izašla na binu]
Marija. ‡ ˆ † Œ ~ encourage that would € ˆ ˆ Ž • on stage
‘Marija is encouraging herself to go out on stage.’

(b) ?Marija hrabri sebe
Marija. ‡ ˆ † encourage self
‘Marija is encouraging herself.’

Since control into purpose clauses is one of the standard diagnostics for the presence of VoiceP (Gehrke 2013, 2014; Alexiadou at al. 2014), one can assume that VoiceP is present with verbs such as the one in (19a) and (20) despite the fact that they reject by-phrases (19a) simply because by-phrases are normally rejected with typical reflexives (21), presumably due to the fact that the agent is already expressed by means of the subject DP/NP.

(21) *Marija se R þ H ã O M D O D strane majke
Marija. ‡ ˆ † Œ ~ comb. ‰ z Œ • from side mother. € ~ ‡
Literally: ‘Marija combed (herself) by her mother.’

When it comes to the licensing of instrumental case-marked DPs/NPs, unlike typical anticausatives, a significant number of psych verb anticausatives allow instrumental case-marked DPs/NPs (22).

(22) (a) Marija se zaprepastila bratovim ponašanjem
Marija. ‡ ˆ † Œ ~ shock. ‰ z Œ • brother's behavior. , ‡ Œ •
'Marija got shocked by/with her brother's behavior.'

(b) Marija se iznenadila njegovim postupkom
Marija. ‡ ˆ † Œ ~ surprise. ‰ z Œ His action., ‡ Œ •
'Marija got surprised by/with his action.'

(c) Marija se oduševila svojim rezultatom
Marija. ‡ ˆ † Œ ~ impress. ‰ z Œ Self's result. , ‡ Œ •
'Marija got impressed by/with her result on the exam.'

Also, unlike typical anticausatives, these constructions tend to reject od('from')-PPs (23).

(23) (a) *Marija se zaprepastila od bratovog ponašanja
Marija. ‡ ˆ † Œ ~ shock. ‰ z Œ • from brother's behavior. € ~ ‡
'Marija got shocked with her brother's behavior.'

(b) *Marija se iznenadila od njegovog postupka
Marija. ‡ ˆ † Œ ~ surprise. ‰ z Œ from his action. € ~ ‡
'Marija got surprised by his action.'

(c) *Marija se oduševila od svog rezultata
Marija. ‡ ˆ † Œ ~ impress. ‰ z Œ from self's result. € ~ ‡
'Marija got impressed by her result on the exam.'

Still, it is crucial to point out that the category of psych verbs is not homogenous when it comes to the licensing of various types of even modifiers. While there are those that allow instrumental NPs/DPs and reject od('from')-PPs, there are also psych-verb anticausatives that exhibit the opposite behavior since they license od('from')-PPs while rejecting instrumental case-marked NPs/DPs (24), but these are significantly less numerous.

(24) (a) Marija se razbesnela *bratovim ponašanjem
Marija. ‡ ˆ † Œ ~ anger. ‰ z Œ • brother. , ‡ Œ • behavior. , ‡ Œ •
/od bratovog ponašanja
from brother's behavior. € ~ ‡
'Marija got angry because of her brother's behavior.'

(b) Marija se U D å D O R V W WXåQ R Ppesmom
Marija. ‡ ˆ † Œ ~ sadden. ‰ z Œ • sad. , ‡ Œ • song. , ‡ Œ •
/od W X å Q Hpesme
from sad. € ~ ‡ song. € ~ ‡
'Marija got sad because of a sad song.'

Finally, there are verbs that can basically tolerate both kinds of event modifiers to varying degrees (25) and native speaker judgments tend to vary with these verbs.

(25) (a) Ivan se iznervirao ?bratovim ponašanjem
Ivan. ‡ ˆ † Œ ~ annoy. ‰ z Œ • brother. , ‡ Œ • behavior. , ‡ Œ •
/?od bratovog ponašanja
from brother's behavior. € ~ ‡
'Ivan got annoyed with/by his brother's behavior.'

(b) Ivan se isprovocirao ?bratovim ponašanjem
Ivan. ‡ ˆ † Œ ~ provoke brother's. , ‡ Œ • behavior. , ‡ Œ •
/?od bratovog ponašanja
from brother's behavior. € ~ ‡
'Ivan got provoked with/by his brother's behavior.'

None of these verbs, however, pattern with pure reflexives because they do not license full reflexive pronouns or control into purpose clauses. The examples in (26a), (26b), and (26c) demonstrate the lack of control into purpose clauses with these verbs while (26a'), (26b'), and (26c') show that they do not license full reflexive pronouns.

(26) (a) *Marija se zaprepastila da bi uplašila Anu
Marija. ‡ ˆ † Œ ~ dazzle. ‰ z Œ • that would scare Ana. z | |
Literally: 'Marija got dazzled in order to scare Ana.'

(a') Marija je zaprepastila sebe
Marija. ‡ ˆ † z Ž ' dazzle. ‰ z Œ • self
Literally: 'Marija dazzled herself'

(b) *Marija se razbesnela da bi uplašila Anu
Marija. ‡ ˆ † Œ ~ anger. ‰ z Œ • that would scare Ana. z | |
Literally: ‘Marija got angry in order to scare Ana.’

(b’) Marija je razbesnela sebe
Marija. ‡ ˆ † z Ž ‘ anger. ‰ z Œ • self
Literally: ‘Marija angered herself’

(c) *Marija se iznervirala da bi uplašila Anu
Marija. ‡ ˆ † Œ ~ annoy. ‰ z Œ • that would scare Ana. z | |
Literally: ‘Marija got annoyed in order to scare Ana.’

(c’) *Marija je iznervirala sebe
Marija. ‡ ˆ † z Ž ‘ annoy.past self
Literally: ‘Marija annoyed herself.’

In sum, among psych verb constructions with SE in Serbian, we clearly find cases that do not pattern with reflexives in all relevant respects, but they cannot be collapsed with typical anticausatives either given the data from event modifier licensing. The “standard approach” to anticausatives seems to be too rigid to account for these cases. Also, semantically, there are strong reasons to believe that the reflexive semantics that Chierchia (2004) proposes for anticausatives fits the denotations of these borderline cases that licensee instrumental case-marked modifiers. Recall that Chierchia (2004) argues that the reason why anticausatives are reflexive is because they denote eventualities that are caused by some internal property or state of the subject. I would suggest that this is precisely what native speakers intend to communicate when they utter a sentence like (27a), and it is precisely this semantic component that makes the difference between (27a) and the corresponding transitive construction (27b). In other words, the two sentences in (27) do not have identical truth-conditional content.

(27) (a) Marija se zaprepastila bratovim ponašanjem
Marija. ‡ ˆ † Œ • dazzle. ‰ z Œ • brother’s behavior. , ‡ Œ •
‘Marija got dazzled by/with her brother’s behavior.’

(b) Bratovo ponašanje je zaprepastilo Mariju
brother’s behavior z Ž ‘ dazzle. ‰ z Œ • Marija. z | |
‘Her brother’s behavior dazzled Marija.’

Even though it might be possible to describe the same situation with either of the two forms, the selection of one instead of the other casts the situation in a different light. The

transitive version (27b) entails that the brother's behavior was the direct cause of the change in Marija's emotional state (it was both a necessary and a sufficient condition) whereas the version with SE communicates that the brother's behavior was a necessary but not sufficient condition for the change in Marija's emotional state and some internal (psychological/emotional) property/state on Marija's part was the decisive factor (see Talmy 1988; Croft 1993; Wolf 2003).

It is difficult to find a test that would pin down this truth conditional distinction, but the contrast in (28) points in that direction quite strongly.

(28) (a) Marija se zaprepastila " " Q L V D SRVHEQR
Marija. ‡ ˆ † Œ ~ dazzle.‰ z Œ Ÿ nothing. , ‡ Œ ¡ special
/ni zbog Þ H J D posebno
‡ ~ € because what special
'Marija got dazzled for no apparent reason.'

(b) *Ništa posebno je zaprepastilo Mariju
nothing special z Ž ' dazzle.‰ z Œ Ÿ Marija. z | |
Literally: 'Nothing special dazzled Marija.'

The (partial) acceptability of (28a) suggests that the construction with SE can be used while explicitly negating the external cause which is impossible with the transitive version (28b). While the prepositional instrumental expression in (28a) is quite degraded, the zbog ('because')-PP introducing a negated indirect cause/reason is perfectly acceptable. What this tells us is that clauses that contain psych-verb anticausatives can receive a truth value and be grammatically acceptable even when the external cause is explicitly negated. On the other hand, when a NP/DP introducing the causer appears as the subject of a transitive sentence (28b), the outcome is ungrammatical. Note that the ungrammaticality of (28b) cannot simply be ascribed to independent factors such as the lack of sentential negation presumably needed to license the Negative Polarity Item (NPI) ništa ('nothing'). If (28b) were ungrammatical because the NPI located in the subject position is not licensed, then, we would expect (28a) to be ungrammatical as well since there is no sentential negation in this example either. Therefore, I assume that the contrast in (28) emerges for semantic reasons. Specifically, the anticausative example (28a) does not entail the existence of an external cause, which is why the external cause can be explicitly negated, but the transitive sentence (28b) includes this entailment, which is why the negation of the existence of the external cause yields a fatal contradiction. It, thus, follows that the two sentences do not have the same truth-conditional content.

4. Modeling the Differences among SE Constructions: Gradients of Reflexivity

To capture the data presented in the previous section, drawing on Sportiche (2014), I propose to treat SE as a variable whose interpretation is structurally determined in two ways: (i) by its merge site, and (ii) the c-commanding NP/DP, which binds it and assigns it interpretation. SE has to be merged with a head within the extended VP domain, which basically means that it can either merge with v^0 or with $Voice^0$ (assuming with Harley 2013 and Merchant 2013 inter alia that these two projections are separate). When it is merged with $Voice^0$, it creates typical agentive reflexives (29a). In that case, it is c-commanded by the subject in Spec VoiceP. The agentive component is, of course, supplied by Voice^0.

(29) (a) Marija se posekla
Marija. ‡ ˆ † Œ ~ cut. ‰ z Œ •
‘Marija cut herself.’

(b) Marija se zaprepastila
Marija. ‡ ˆ † Œ ~ dazzle. ‰ z Œ •
‘Marija got dazzled.’

(c) Stolica se polomila
chair. ‡ ˆ † Œ ~ break. ‰ z Œ •
‘The chair broke.’

SE can also be merged with v^0 in the absence of VoiceP, and in that case, two things can happen. It can either be merged with v^0 that projects a Spec position/an external argument or it can merge with v^0 without the Spec position. When there is an NP/DP in the Spec vP position, this NP/DP will bind the SE variable and assign it reference giving rise to a kind of “semi reflexive” denotation that we can see with psych verb anticausatives (29b). In such cases, the absence of VoiceP will ensure the lack of agentivity and the external argument will be interpreted as “self-cause” rather than “self-agent” yielding the crucial difference between these constructions and typical reflexives.

Alternatively, SE can merge with Spec-less v^0, in which case there would be no NP/DP c-commanding it and it would receive arbitrary interpretation as with typical anticausatives (29c). Finally, in order for this analysis to work, the binding domain for SE would have to be restricted to the first-phase or the extended VP because, otherwise, it would be bound by the subject NP/DP once it reaches Spec TP even with typical anticausatives.

To illustrate the proposal more clearly, I submit the following structural representations for pure reflexives, semi-reflexives and pure anticausatives in (30). (30a) illustrates a reflexive structure where SE merges with $Voice^0$ with the agent NP/DP in the

Spec position. (30b), on the other hand, represents the new category of “semi-reflexives” where SE merges with v^0, but the VoiceP layer is not projected. Nevertheless, the Spec vP position is still available for the external argument NP/DP denoting causer. Finally, the “typical anticausative” construction represented by (30c) does not project the VoiceP layer, and it lacks the Spec vP position for the external argument. Consequently, there is no hierarchically higher NP/DP, within the domain of the “first phase” / extended VP to bind the variable introduced by SE. This lack of binding does not result in ungrammaticality because the variable can receive arbitrary interpretation.

(30) D U H À H [L Y H

E V H P L U H À H [L Y H

(c) anticausative

As a side note, the idea that SE is a variable whose interpretation is structurally determined is also desirable with respect to middles and impersonal constructions where it gets arbitrary interpretation. The impersonal construction in (31a) comes with the

SE morpheme and the external argument of the verb (the agent) is left unspecified. Similarly, with middles (31b), the external argument or the agent is also unspecified, and, again, one finds the morpheme SE in these constructions as well.

(31) (a) U biblioteci se þ X Y D knjige
in library. … ˆ | Œ ~ keep books. ‡ ˆ †
'The library is where books are kept.'

(b) Ova knjiga se lako þ L W D
this book. ‡ ˆ † Œ ~ easily read
'This book reads easily.'

However, providing a precise analysis of these constructions goes beyond the scope of this paper.

The different structural representations in (30) now enable us to explain the contrasts L Q W K H O L F H Q V L Q J R I G L I I H U H Q W W \ S H V R I H Y H Q W those psych-verb anticausatives that license instrumental NPs/DPs would correspond to "semi-reflexive" configurations (30b). Assuming that instrumental NPs/DPs are licensed by the presence of an external argument either in Spec vP or Spec VoiceP, we would have an account of the acceptability of these elements with agentive transitives, typical reflexives (30a) and structures we have called "semi-reflexives" (30b). The unacceptability of instrumental NPs/DPs with typical anticausatives (30c) as well as unaccusatives and the use of od('from')-PPs would, then, follow from the lack of the external argument in Spec vP.

Finally, I would like to point out that the present approach has the potential to shed some light on the puzzling lack of psych verb anticausatives in English. Namely, because there is no SE in this language, it is impossible to generate the kind of semi-reflexive construction that we find in Serbian (and presumably other languages that have this morpheme). Merging the experiencer DP in Spec vP position as in (30b) in the absence of SE would leave the structure without the internal argument resulting in ungrammaticality.

In terms of broader implications, this account bridges the gap between the "reflexive account" and "standard account" of SE by suggesting that reflexivity is a graded category. Contrary to the "standard account", it avoids the severance of all syntactic or semantic links between the reflexive and anticausative uses of SE. At the same time, in maintaining this link, it does not go so far as to extend the same reflexive semantics to all anticaustive structures. Instead, the present account confines the semantic connection between reflexives and anticausative uses of SE to the fact that SE introduces a variable, which can receive the interpretation from different NPs depending on the structural configuration it finds itself in. With typical reflexives,

SE receives its interpretation from the agent argument in Spec VoiceP; with "semi-reflexives", its interpretation is determined by the cause argument in Spec vP; finally, in typical anticausative constructions (potentially also in middles and impersonals), SE remains unbound receiving arbitrary interpretation.

The account also raises the question of the syntactic and semantic status of SE forms with psych verbs in other languages that have this morpheme. We have seen that these constructions combine with from-PPs expressing causers in Romanian, but in Greek they combine with[illegible] with 3 3 V W \ S L F D O R I L Q V W U X P H Q W V $ O H [L D G R issue whether these facts signal different syntactic and semantic behaviors of SE forms in these languages is a topic for further research.

5. Conclusion

This paper addressed the broader issue of the status of anticausatives by focusing on the event modifier licensing with Serbian psych verb anticausatives. It was established that like reflexives, but unlike typical anticausatives, psych verb anticausatives in Serbian can license instrumental case-marked expressions of cause and reject od('from')-PPs. Still, these constructions are not fully reflexive as they disallow control into purpose clauses and the full reflexive pronoun sebe ('self').

This data was interpreted in light of Sportiche's (2014) approach to SE arguing that Chierchia's (2004) "reflexive approach" to anticausatives as well as the so-called 'standard view' (Schäfer and Vivanco 2016), which denies any kind of synchronic link between reflexive and anti-causative SE, are simply too rigid to explain these data. I argued that what is needed is a graded or ranked model of reflexivity constrained by the options provided by the extended VP structure, which allows for a category between pure reflexives and pure anticausatives. I called this category "semi-reflexives".

The observations were modeled by assuming that SE introduces a variable that has to be bound by a c-commanding NP/DP in order to obtain interpretation. With pure reflexives, SE attaches to Voice and gets bound by the agent in Spec VoiceP. With "semi-reflexives", SE combines with v and gets bound by the causer in Spec vP. Finally, with anticausatives, SE is merged with v that does not project Spec, a position, which is why it remains unbound and receives arbitrary interpretation. Crucially, the binding domain for SE has to be restricted to the first phase (extended VP). The distribution of instrumental NPs/DPs and od('from')-PPs with different types of constructions involving SE was accounted by assuming that instrumental NPs/DPs need to be licensed by the presence of an external argument in Spec vP or Spec VoiceP. This explains why these expressions are licensed with agentives, reflexives and "semi-reflexive" psych-verb anticausatives but banned with typical anticausatives and unaccusatives, where the only option for expressing the cause argument is the od('from')-PP.

Works Cited

Alexiadou, Artemis, and Elena Anagnostopoulou. 2009. “Agent, Causer and Instrument PPs in Greek.” MIT Working Papers in Linguistics í

$OH[LDGRX $UWHPLV DQG *LDQQLD ,RUG FKLRDLD Lingua í

Alexiadou, Artemis, Elena Anagnostopoulou, and Florian Schäfer. 2015. External Arguments in Transitivity Alternations: A Layering Approach. Oxford: Oxford University Press.

Alexiadou, Artemis, Berit Gehrke, and Florian Schäfer. 2014. “The Argument Structure of Adjectival Participles Revisited.” Lingua í

$OH[LDGRX $UWHPLV *LDQQD ,RUG FKLRDLD 0DU and Florian Schäfer. 2013. “The Realization of External Arguments in Nominalizations.” The Journal of Comparative Germanic Linguistics í í

%HOOHWWL $GULDQD DQG /XJL 5L]]L Natural Language & Linguistic Theory í

Chierchia, Genaro. 2004. “A Semantics for Unaccusatives and Its Syntactic Consequences.” In The Unaccusativity Puzzle: Exploration of the Syntax-Lexicon Interface, edited by Artemis Alexiadou, Eelena Anagnostopoulou, and Martin Everaert, í 2[IRUG 2[IRUG 8QLYHUVLW\ 3UHVV

Croft, William. 1993. “Case Marking and the Semantics of Mental Verbs.” In Semantics and the Lexicon, HGLWHG E\ -DPHV 3XVWHMRYVN\ í

Gehrke, Berit. 2013. “Still Puzzled by Adjectival Passives.” In Syntax and Its Limits, HGLWHG E\ 5DIIDHOOD)ROOL &KULVWLQD 6HYGDOL Oxford University Press.

Gehrke, Berit. 2015. “Adjectival Participles, Event Kind Modification and Pseudo-Incorporation.” Natural Language & Linguistic Theory í

Gehrke, Berit, and Cristina Marco. 2014. “Different By-Phrase with Adjectival and Verbal Passives: Evidence from Spanish Corpus Data.” Lingua í

Harley, Heidi. 2013. “External Arguments and the Mirror Principle: On the Distinctness of Voice and v.” Lingua í

Koontz-Garboden, Andrew. 2009. “Anticausativization.” Natural Language and Linguistic Theory í

Koontz-Garboden, Andrew. 2012. “The Monotonicity Hypothesis.” In Telicity, Change DQG 6WDWH $ &URVV í &DWHJRULDO 9LHZ RI (YHQW DQG 9LROHWD 'HPRQWH í 2[IRUG 2[IRUG 8QL

Kratzer, Angelika. 1994. 7KH (YHQW $UJXPHQW DQG WKH 6HPDQWLFV of Massechusetts.

Landau, Idan. 2009. The Locative Syntax of Experiencers. Cambridge: MIT Press.

Levin, Beth. 1993. (QJOLVK 9HUE &ODVVHV DQG $OWHUQDWLRQV University of Chicago Press.

Merchant, Jason. 2013. “Voice and Ellipsis.” Linguistic Inquiry í
Parsons, Terence. 1990. Events in the Semantics of English. Cambridge: MIT Press.
Pesetsky, David. 1996. Zero Syntax: Experiencers and Cascades. Cambridge: MIT Press.
Schäfer, Florian, and Margot Vivanco. 2016. “Anticausatives Are Weak Scalar Expressions, Not Reflexive Expressions.” Glossa: A Journal of General Linguistics 1 (1).
Sportiche, Dominique. 2014. “French Reflexive se: Binding and Merge Locality.” In Locality, edited by Enoch Olade Aboh, Maria Teresa Guasti, and Ian Roberts, í 2 [I R U G 2 [I R U G 8 Q L Y H U V L W \ 3 U H V V
Talmy, Leonard. 1988. “Force Dynamics in Langauge and Cognition.” Cognitive Science 37 í
Wolff, Phillip. 2003. “Direct Causation in the Linguistic Coding and Individuation of Causal Events.” Cognition í

The Syntax of Locative Inversion in Mandarin Chinese

Chang Liu

Université Paris 8/UMR 7023 CNRS, Paris, France;
Université de Picardie Jules Vernes, Amiens, France

changliu1120@gmail.com

Abstract: This paper examines the verbs that can participate in the formation of the Locative Inversion with an obligatory localiser phrase in Mandarin Chinese. We show evidence that runs counter to the previous observation that the Locative Inversion with the perfective aspect -le has an implicit Agent, as opposed to the Locative Inversion with the imperfective/durative -zhe. We argue that whether there is an implicit Agent is dependent on the verbal argument structure, not directly on the aspectual differences. Based on the interpretive differences induced by the occurrence of the aspectual markers, we provide a non-uniform analysis of the Locative Inversion based on the verb classes identified in this paper.

Keywords: Locative Inversion; Aspect; Syntax; Mandarin Chinese

1. Introduction

This paper investigates the syntax of Locative Inversion (Henceforth, LI) in Mandarin Chinese. This construction has the surface structure "localiser phrase + V + Asp + noun phrases" (Fan 1963; Huang 1987; Nie 1989; Hu 1995; Pan 1998; Yang and Pan 2001; F. Liu 2007; Lin 2008; Paul et al. 2019; Zhang 2019). As shown in (1), the localiser phrase F Ì Q 'village' precedes the verb V ƒ 'die', and the post-verbal Theme/Patient is \ t J H U p 'one person'. In (2), the localiser phrase T L i Q J V K j Q J 'wall-on' precedes the verb J 'hang', which can be suffixed by either the perfective marker -le, or the durative/imperfective marker -zhe. The post-verbal Theme argument can be a definite demonstrative phrase Q j I ~ K X j 'that painting', in contrast with the indefinite Theme/patient in (1).

(1) F Ì Q O ÿ V ÿ -le yí-ge rén
village-in die - ‰ ~ ‹ • one- | . person
“In the village died one person.”

(2) qiáng-shàng guà -zhe/-le nà-fú huà
wall-on hang - } Ž ‹/- ‰ ~ ‹ • that- | … • painting
Lit. “On the wall is/was hung that painting.”

We will examine the LI based on the verbs like X j ‘hang’. This paper is organised as follows. Section 2 presents the background of LI in Mandarin Chinese. Section 3 presents the verb classes that can participate in the formation of LI. Section 4 shows the distinction between the LI with the perfective -le and the LI with the durative/imperfective -zhe. They differ in whether there is an implicit Agent and in the complexity of verbal argument structure. We further show that with -zhe, the LI can be interpreted with a state reading, a progressive reading or both, depending on the verbs. In Section 5 we provide a non-uniform analysis. We conclude in Section 6.

2. Background

This construction has been referred to as existential sentences (Huang 1987), presentative sentences (Hu 1995), existential structure (Yang and Pan 2001) and locative construction (Paul et al. 2019) in the literature. Huang (1987) divided the verbs that can form existential sentences (or / ˆ | z • , • ~ , ‡ • ~ ‹ Œ , ˆ ‡ in our term) into two classes. Huang regard (3) based on the verbs like V ƒ ‘die’ as Type II existential sentences, and (4) as Type III existential sentences.

(3) F Ì Q O ÿ V ÿ -le yí-ge rén 7 ’ ‰ ~ , ,
village-in die - ‰ ~ ‹ • one- | . person
“In the village died one person .”

(4) *(qiáng-shàng) guà -zhe/-le nà-fú huà 7 ’ ‰ ~ , , ,
wall-on hang - } Ž ‹/- ‰ ~ ‹ • that- | … • painting
Lit. “On the wall is/was hung that painting.”

Huang observed that the localiser phrase is optional in Type II, but obligatory in III Type and that the post-verbal noun phrase must be indefinite in Type II, but does not have to be indefinite in Type III. The obligatory presence of the localiser phase in Type III suggests that it has an argument status. In this paper, we will not analyse the verbs involved in Type II, which are basically unaccusative verbs.

(5)		Localiser phrase	Definiteness Restriction
	Type II	optional	yes
	Type III	obligatory	no

The localiser phrase in Mandarin Chinese has the distribution of DP argument (Huang, Li, Li 2009). As shown in (6), the localiser phrase can occur in subject or object position.

(6) (a) > W — P H Q G H /[kéng-wài] K@ K Q P L O u
their city /city-outside very beautiful
³ > 7 K H L U F L W \ @ > W K H R X W V L G H R I W K H F L W \ @ L

(b) Z qù-guò > W — P H Q G H /[kéng-wài] K@
I go- ~ ‘ ‰ their city / city-outside
³ , K D Y H E H Q W R > W K H L U F L W \ @ > W K H R X W V L G H

Furthermore, the literature has been particularly interested in the aspectual markings on the LI based on the verbs of Huang’s Type III. In (4), the verb J X j ‘hang’ can be suffixed by either a perfective aspect -le or a durative/imperfective -zhe. While some scholars considered these two markers freely interchangeable (Fan 1963, Nie 1989), others claimed that they are not semantically identical (Hu 1995, Yang and Pan 2001, F. Liu 2007), and further observed that the LI with the perfective -le has an implicit agent, whereas the LI with the durative -zhe does not have an implicit agent (Pan 1996, Paul et al. 2019).

In the next section, we present the verbs that can form the Locative Inversion with an obligatory localiser phrase. Based on the differences among the subclasses, we argue that it is the verbal argument structure that determines the presence of an Agent, not the aspectual markers.

3. Verbs

We identify five subclasses of verbs that can form the Locative Inversion with an obligatory localiser phrase (7). These verbs include but are limited to the locational verbs of Huang’s (1987) Type III existential sentences. We will present them based on (i) whether they can be suffixed by the perfective -le or the durative -zhe, and (ii) the interpretations resulted from the possible aspectual markings.

With the imperfective/durative -zhe, the LI based on these verbs can have a stative reading, or a progressive reading, or both. The possible readings are related to the two uses of -zhe in Smith (1997). In its basic use, zhe focuses on the state of position and posture, or the state that results from an event. In its extended use, -zhe ‘presents internal stages of durative events as static’ (Smith 1997; 274).

(7) hang-verbs:] K j Q ‘stand’,] X z ‘sit’, G Ì Q ‘squat on heels’, J X u ‘kneel on knees’, W
‘lie on back’, S — ‘lie on stomach’, I j Q J ‘put’, J X j ‘hang’; N j R ‘lean’, F K X — Q ‘w
Z p L ‘surround’, G Q J ‘wait’, J j L ‘cover’
write-verbs: [L‘write’, K X j ‘draw’, kè ‘carve’, O L ~ ‘leave/remain’, O V K • Q ‘J
‘raise (a flag)’, M u ‘record, document’
walk-verbs:] … X ‘walk’, S • R ‘run’, I • L ‘fly’, pá ‘crawl/creep’, \ y X ‘swim’, W L j
‘jump’, J ‡ Q ‘roll’, S L — R ‘float’
plant-verbs:] K z Q J ‘plant’, M L j Q ‘build, construct’, J j L ‘build’, S Ì ‘lay, pave’, W
‘stick’
sing-verbs: F K j Q J ‘sing’, N — L ‘hold (me] K g s),
‘book’, P j R ‘em R’ j L ‘sell’

+ D Q J-verbs correspond to Levin and Rappaport Hovav’s (1995) • ~ ‹ { Œ ˆ • Œ %
| ˆ ‡ • , € Ž ‹ z • , ˆ ‡ and • ~ ‹ { Œ ˆ • ‰ Ž • • , ‡ € ˆ ‹ ‰ Ž • • , ‡ € , ‡ z Œ ‰ z • , z .
sec. 6.4.5). As described by the LI based on these verbs, the entity participant
spatial configuration with a certain position. As shown in (8), they can be suffixed either
by the durative -zhe or by the perfective (Pan 1996; Liu 2007). With -zhe, the sentence
describes a state. The state reading can be diagnosed by the incompatibility with the
progressive] j L-, which rejects states (9).

(8) (a) P p Q N Xzhàn -zhe/-le M ÿ J H O ý R V K ¯
doorway stand - } Ž ‹/- ‰ ~ ‹ a few- | … • teacher
“At the door stand a few teachers.”

(b) qiáng-shàng guà -zhe/-le \ u I ~ huà
wall-on hang - } Ž ‹/- ‰ ~ ‹ one- | … • painting
Lit. “On the wall was hung one painting.”

(9) (a) P p Q N Xzài zhàn -zhe M ÿ J H O ý R V K ¯
doorway ‰ ‹ ˆ € stand - } Ž ‹ a few- | … • teacher
(“At the door were standing a few teachers.”)

(b) *qiáng-shàng zài guà -zhe \ u I ~ huà
wall-on ‰ ‹ ˆ € hang - } Ž ‹ one- | . painting
(“On the wall was being hung one painting.”)

: U L W H-verbs correspond partially to Levin and Rappaport Hovav’s (1995: sec 6.4.5)
• ~ ‹ { Œ ˆ • , † z € ~ , † ‰ ‹ ~ Œ Œ , ˆ ‡. They can be suffixed either by the durative -zh
perfective -le. With -zhe, the sentence describes a state, which is incompatible with the
progressive] j L-.

(10) (a)] K X À] L V K j [L J -zhe/-le yí-ge] u
table-on write - } Ž ‹/- ‰ ~ ‹ • one- | … • character
“On the table was/is written one character.”

(b)] K X À] L V k à i Q J [L -zhe yí-ge] u
table-on ‰ ‹ ˆ € write - } Ž ‹ one- | … • character
(“On the table was being written one character.”)

Walk-verbs correspond to Levin and Rappaport Hovav’s (1995) • ~ ‹ { Œ Ÿ ˆ • † z ‡ ‡ ~ ‹ † ˆ • , ˆ. ‡ Contrasting with the previous two groups, walk-verbs can only be suffixed by the durative -zhe, not by the perfective -le (1a–b, cf. Yu 1995). In addition, with -zhe, the LI based on walk-verbs describes an ongoing dynamic event (Nie 1989), and is compatible with the progressive] j L- (11b). The incompatibility with the perfective -le is due to the fact that the LI based on walk-verbs describes unbounded activities, which cannot satisfy the bounded requirement imposed by the perfective (see Xiao and McEnery 2004 on -le).

(11) (a) F — R F K ý Q J] V K j Q J -zhe/*-le [G X À [X p V K • Q J
playground-on walk - } Ž ‹/- ‰ ~ ‹ • many student
Lit. “On the playground are walking many students.”
(NO: “on the playground walked many students.”)

(b) F — R F K ý Q J z à i K j Q J X -zhe [G X À [X p V K • Q J
playground-on ‰ ‹ ˆ € walk - } Ž ‹ many student
Lit. “On the playground are walking many students.”

3 O D Q W-verbs can be suffixed either by the durative -zhe or by the perfective -le. With the durative -zhe, the LI based on S O D Q W-verbs is ambiguous between a state reading and a progressive reading (12a, cf. Yeh 1993; Smith 1997). The insertion of the progressive] j L- in (12b) can single out the progressive reading, while excluding the state reading. The LI with the perfective -le in (12c) has a resultant state reading.

(12) (a) \ X i Q] L O z h ò n g -zhe \ t N • shù.
yard-in plant - } Ž ‹ one- | … • tree
(i) “In the yard is planted one tree.” —> state reading
(ii) “In the yard is being planted one tree.” —> progressive reading

(b) \ X i Q] L Zài zhòng -zhe \ t N • shù.
yard-in ‰ ‹ ˆ € plant - } Ž ‹ one- | … • tree
~~(i) "In the yard is planted one tree." —> state reading~~
(ii) "In the yard is being planted one tree." —> progressive reading

(c) \ X i Q] L Zhòng -le \ t N • shù.
yard-in plant - ‰ ~ ‹ • one- | … • tree
"In the yard was planted one tree."

6 L Q J-verbs (cf. Fan 1963) can be suffixed either by the durative -zhe or by the perfective -le. With the durative -zhe, the LI describes an on-going dynamic event, which is then compatible with the occurrence of the progressive] j L- (13b). With the perfective -le, the LI describes a terminated event without the implication of a resultant state.

(13) (a) tái-shàng chàng -zhe/-le J • M •
stage-on sing - } Ž ‹/- ‰ ~ ‹ • opera
-zhe: "On the stage is being sung the opera."
-le: "On the stage was sung the opera."

(b) tái-shàng zài chàng -zhe J • M •
stage-on ‰ ‹ ˆ € sing - } Ž ‹ opera
"On the stage is being sung the opera."

We resume the five subclasses in (14). With the durative -zhe, the LI based on K D Q J-verbs and Z U L W H-verbs describes states, the LI based on walk-verbs and V L Q J-verbs describes on-going dynamic events, and the LI based on S O D Q W-verbs is ambiguous between the two readings. With the the perfective -le, the LI based on V W D Z U L W H-verbs, Z U L W H-verbs and S O D Q W-verbs describes resultant states, while the LI based V L Q J-verbs describes a terminated event. The LI based on walk-verbs is incompatible with the perfective -le.

(14)		durative -zhe		perfective -le
		state	progressive	
(a)	K D Q J-verbs	yes	no	yes
(b)	Z U L W H-verbs	yes	no	yes
(c)	walk-verbs	no	yes	*
(d)	S O D Q W-verbs	yes	yes	yes
(e)	V L Q J-verbs	no	yes	yes

4. Perfective -le vs Imperfective -zhe in Locative Inversion

The literature on Mandarin LI has paid special attention to the contrast between the perfective and imperfective LI: diagnostics show that the LI with the perfective -le has an implicit agent, whereas the LI with the durative -zhe does not have an implicit agent.

In section 4.1, we will show evidence that runs counter to this claim. With the perfective -le, while the LI based on J X j 'hang'/ [L 'write'-verbs and] K z Q J 'plant'-verbs can be diagnosed with an implicit agent, the LI based on F K j Q J 'sing'-verbs cannot. We argue that whether there is an implicit Agent is determined by the verbal argument structure.

In section 4.2, we use the \ z X 'again'-test to show that, with the perfective -le, the LI based on J X j µ K D Q J ¶ [L µ Z] K z Q J 'plant'-verbs has a result component and a state component, whereas the LI based F K j Q J 'sing'-verbs only has an event component. In addition, with the durative -zhe, the LI based on J X j['hang'/ 'write'-verbs has a state reading without event implication, while the LI based on F K j Q J 'sing'-verbs and] … X 'walk'-verbs has a progressive reading.

4.1 Implicit Agent

As observed in H. Pan (1996), F. Liu (2007) and Paul et al. (2019), with the-perfective -le, the LI has an implicit agent, whereas the LI with the durative -zhe does not have an implicit agent. These tests, which diagnose the existence of an implicit Agent, include the occurrence of the passive morpheme (Pan 1996), the subject-oriented adverbial J • \ u 'deliberately' (F. Liu 2007) and the purpose clause (see also Paul et al. 2019). As shown in (15), only the LI with -le can occur with these elements, whereas the LI with -zhe cannot.

(15) (a) With the passive E q L
qiáng-shàng bèi = K — Q J guà *-zhe/-le \ u I ~ huà.
wall-on ‰ z Œ Zhangsan hang - } Ž ‹/- ‰ ~ ‹ one- | … • painting
Lit. "On the wall was hung one painting by Zhangsan."

(b) With a subject-oriented adverbial 'deliberately'
qiáng-shàng gùyì guà *-zhe/-le \ u I ~ huà.
wall-on deliberately hang } Ž ‹ ‰ ~ ‹ one- | … • painting
Lit. "On the wall was hung one painting deliberately."

(c) With a purpose clause

wèile \ t Q J M L ⊙ ý R V K ¯ qiáng-shàng guà
in.order.to greet teacher wall-on hang
*-zhe/-le \ u W L i R K X — Q \ t Q J L — R \
- } Ž ‹/- ‰ ~ ‹ one- | … • welcome banner
“In order to greet the teachers, on the wall is hung one welcome banner.”

Why is the presence of implicit Agent in LI dependent on the perfective -le? This question has not been answered. However, as shown in (16), the LI based on F K j Q J ‘sing’- runs counter to the correlation between the presence of an implicit Agent and the perfective - O H contrast with J X j ‘hang’-verbs as in (15), the LI based F K j Q J ‘sing’-verbs is not compatible with the passive bèi or subject-oriented adverbial, even with the perfective -le.

(16) (a) *tái-shàng bèi = K — Q J V chàng -le J • M •
stage-on ‰ z Œ Ž Œ Zhangsan sing - ‰ ~ ‹ • opera
(Lit. “On the stage was sung the opera by Zhangsan.”)

(b) *tái-shàng gùyì chàng -le J • M •
stage-on deliberately sing ‰ ~ ‹ • opera
(Lit. “On the stage was sung the opera deliberately.”)

Furthermore, we cast doubt on the test with the passive morpheme bèi. As shown in (17), the passive morpheme bèi is in fact obligatory in the formation of canonical passives in Mandarin Chinese. However, the LI in (15a) shows that this morpheme is only optional in the LI with -le. The passive morpheme bèi is not a counterpart of English passive by (Li 1990). Hence, the LI with the perfective -le must be distinguished from the canonical passives.

(17) (a) = K — Q J V *(bèi) O ÿ V G ý -le.
Zhangsan ‰ z Œ Ž Œ Lisi hit - ‰ ~ ‹ •
“Zhangsan was hit by Lisi.”

(b) = K — Q J V *(bèi) G ý -le.
Zhangsan ‰ z Œ Œ hit - ‰ ~ ‹ •
“Zhangsan was beaten.” + X D Q J / L D Q G / L

In this subsection, we have shown evidence that runs counter to the previous observation that the perfective LI has an implicit Agent. As shown in (18), in contrast with J X j ‘hang’/ [L ‘write’/] K z Q J ‘plant’-verbs, the perfective LI based on F K j Q J ‘sing’-verbs

does not have an implicit Agent. We will argue in section 5.2 that the occurrence of an implicit Agent is determined by the argument structure in which verbs can occur.

(18) Is there an implicit Agent in LI?

	} Ž ‹ z • , • ~	“ • ~perfective -le
(a)	K D Q J-verbs no	yes
(b)	Z U L W H-verbs no	yes
(c)	walk-verbs no	*
(d)	S O D Q W-verbs no	yes
(e)	V L Q J-verbs no	no

4.2 Structural Complexity

Based on the \ z X ‘again’-test, we bring about more contrast between the LI with the perfective -le and the LI with the durative -zhe.

We begin with the LI with the perfective -le. The \ z X ‘again’-test further sets F K j Q J ‘sing’-verbs apart from other verbs as exemplified by J X j ‘hang’. As shown in (19), with \ z X ‘again’, the LI based on J X j ‘hang’ is ambiguous. In the reading (i), the sentence presupposes that there was a previous state of one painting hung on the wall. In the reading (ii), the sentence presupposes a previous event of hanging. Von Stechow (1996) has used the German Z L H G H U ‘again’ to diagnose the presence of two eventualities in change-of-state predicates: a dynamic event and a state. Similarly, the two readings in (19) can be explained in the same way. In the reading (i) \ z X ‘again’ is associated with the state component of the structure involved in LI, while the \ z X ‘again’ scopes over the dynamic event of ‘hanging one painting’ in the reading (ii).

(19) qiáng-shàng yòu guà -le \ u I ~ huà.
wall-on again hang ‰ ~ ‹ • one- | … • painting
L ³ 2 Q W K H Z D O O W K H V W D W H R I íRQH-HV WDLQWLQJ-J H
L L ³ 2 Q W K H Z D O O D Q R W K H U S D L Q W L Q J Z L W K Q J

By contrast, the LI based on F K j Q J ‘sing’-verbs does not have a state component. As shown in (20), there is only one dynamic event of ‘singing opera’ which can be associated with the \ z X ‘again’.

(20) tái-shàng yòu chàng -le J • M •
stage-on again sing ‰ ~ ‹ • opera
“On the stage the event of singing the opera happened” again.

The LI with the durative -zhe shows different results than the the LI with the perfective -le. As shown in (21), the adverb \ z X ‘again’ can only target a state component in the LI with -zhe based on J X j ‘hang’. The sentence presupposes a previous state of one painting hung on the wall. Unlike (19), (21) does not have a repetitive reading, suggesting that there is no event implication in the structure.

(21) qiáng-shàng yòu guà -zhe \ u I ~ huà.
wall-on again hang } Ž ‹ one- | … • painting
L ³ 2 Q W K H Z D O O W K H V W D W H R I í RQH-V SDLQWLQJ-H K
(ii) ~~“On the wall another painting was hung.”~~ ~~í repetitive~~

By contrast, with the durative -zhe, the LI based on F K j Q J ‘sing’-verbs (and] … X ‘walk’-verbs) has a progressive reading. The sentence (22) with \ z X ‘again’ presupposes a previous on-going event of ‘singing the opera’. It indicates that \ z X ‘again’ scopes over the dynamic event.

(22) tái-shàng yòu chàng -zhe J • M •
stage-on again hang } Ž ‹ opera
“On the stage there is again an event of singing the opera.

Recall that, with the durative -zhe, the LI based on] K z Q J ‘plant’ is ambiguous between a state reading and a progressive reading. When the adverb \ z X ‘again’ is inserted as in (23), the sentence can (i) presuppose a previous state of one tree planted in the yard (like 19i), or (ii) presuppose a previous on-going event of planting one tree (like 22).

(23) \ X i Q] L O ï yòu zhòng -zhe \ u N • shù.
yard-in again plant } Ž ‹ one- | … • tree
(i) “In the yard the state of one tree planted in the yard occurs again.”
(ii) “In the yard there is again an event of planting one tree.”

To summarise, we use the \ z X ‘again’-test to show that the LI is not structurally X Q L I R U P L Q W K H / , F D Q K D Y H D F R P S O H[ ] G \ Q D P ‘hang’/ [L ‘write’/] K z Q J ‘plant’-verbs, but a dynamic event structure with F K j Q J ‘sing’/] … X ‘walk’-verbs. The state in the complex structure results from the dynamic event, that is, a resultant state. The resultant state in the LI with the perfective -le must be distinguished from the state without event implication in the LI with the durative -zhe. In the following sections, we will explain why the perfective -le can occur with the complex structure, and why the durative -zhe somehow ‘forces’ a simple structure.

(24) Structural composition

	durative -zhe	perfective -le
(a) K D Q J-verbs	state	dynamic event + state
(b) Z U L W H-verbs	state	dynamic event + state
(c) walk-verbs	dynamic event	*
(d) S O D Q W-verbs	dynamic event or state	dynamic event + state
(e) V L Q J-verbs	dynamic event	dynamic event

5. Analysis

5.1 Base Structure

(25) shows the transitive sentences without Agent. The localiser phrase can either follow the theme argument \ u N • V K • ‘one tree’, or precede the verb] K z Q J ‘pl However, while the perfective -le can occur in these two sentences, the durative/imperfective -zhe can only occur in (25b) in which the localiser phrase precedes the verb. We take the contrast with respect to the possibility of -zhe to indicate that these two sentences have different structures.

(25) (a) = K — Q J V zhòng -*zhe/-le \ t N • shù zài \ X i Q] L O ÿ
Zhangsan plant - } Ž ‹/- ‰ ~ ‹ one- | … • tree ‰ ‹ ~ ‰ yard-in
“Zhangsan planted one tree in the yard.”

(b) = K — Q J V zài — Q \ X i Q] L zhòng ÿ -zhe/-le \ t N • shù.
Zhangsan ‰ ‹ ~ ‰ yard-in plant } Ž ‹ ‰ ~ one- | … • tree
perfective -le: “Zhangsan planted one tree in the yard.”
durative -zhe: “Zhangsan is planting one tree in the yard.”

Why is the durative -zhe unable to occur in (25a)? The structure proposed as in (26) accounts for it. In (26), the sequence following the verb] K z Q J ‘plant’ is a locational structure pP in which the localiser phrase \ X i Q] L O is in complement of P, and the Theme \ u N • V K • ‘one tree’ is introduced by Theme P. The pP denotes states. The verb] K z Q J ‘plant’, which is formed by the root being adjoined to the eventive verbaliser v, takes the state-denoting pP as complement. The vP has a change-of-state meaning: the event of planting results in the state of one tree being in the yard. Above the vP, an aspectual marker is merged in Asp2 (Tsai 2008, Travis 2004). Since the durative -zhe selects an atelic predicate (Jo-wang Lin 2002), the telic change-of-state vP is incompatible with it. Above AspP2, the Agent = K D Q J V D Q is introduced by Kratzer’s (1996) Voice.

(26)

The little p can be regarded as the prepositional counterpart of little v or Voice. The little p introduces the Figure argument \ u N • V K • ‘one tree’, and assigns or values th Case of the Ground argument \ X i Q] L O ƒ ‘yard-in’ (Svenonius 2010, Levinson 2011 reminiscent of little v or Voice which introduces the external argument and assigns the Accusative Case to the internal argument (Burzio’s generalisation). Note that we assume that] j t at’ is merged in P (see also A. Williams 2016). Later, we will take (26) as the base structure for the perfective LI based on ‘hang/write/plant’-verbs. Since (26) excludes the durative -zhe, it explains why -zhe cannot occur in this type of complex [dynamic H Y H Q W V W D W H @ V W U X F W X U H F I

5.2 Locative Inversion with an Implicit Agent

Recall that the perfective LI based on J n X j ‘hang’-verbs [, L ‘write’-verbs and] d K z Q J ‘plant’ verbs is diagnosed with an implicit Agent. This type of LI must be distinguished from the canonical passive structure with bèi µ S D V V ¶ : H H [W H Q G W K H > of (26) to the perfective LI with an implicit Agent in (27a). As shown in (27b), the

localiser phrase \ X i Q] L O ƒ ‘yard-in’ is in complement of the null P, while the Theme \ u V K • ‘one tree’ is introduced by the little

(27) (a) \ X i Q] L Ozyhòng -le \ t N • shù.
yard-in plant - ‰ ~ ‹ • one- | … • tree
Lit. “In the yard was planted one tree (by someone).”

(b)

We propose that the implicit Agent is introduced by a different type of Voice, which introduces an external argument variable, not an overt external argument in its specifier. This type of head is similar to the Middle Voice proposed for Greek by Alexiadou, Anagnostopoulou and Schäfer (2015), Alexiadou and Doron (2012).

The localiser phrase originates within vP (pace Paul et al. 2019), and moves to Spec,TP (pace Yu 1995).[1] We further assume that Voice$_{middle}$( ) and little p are phase heads, and the localiser phrase \ X i Q] L O ‘yard-in’ moves successive-cyclically through the edge of pP and VoiceP to avoid being trapped inside the lower phases, cf. Chomsky’s (2000) 3 • z Œ ~ , † ‰ ~ ‡ ~ • ‹ z { , … , • ’ & ˆ ‡ } , • , ˆ ‡. The movement can be triggered by the E feature in pP and VoiceP, and the EPP in T.

Note that the movement of the localiser phrase \ X i Q] L O ‘yard-in’ should have been intervened or blocked by the Theme \ u N • V K • ‘one tree’ in Spec,pP. Assumin Chomsky’s (1995) (Š Ž , } , Œ • z ‡ | ~ thanks to the movement of P to p, the edge of p (Spec,pP) and the Theme in Spec,pP are equidistant from the localiser phrase in the complement of P, because the edge of pP and Spec,pP are now in the same minimal domain. Another solution is … ~ z ‰ • ‹ (Bobaljik 1995, Legate 2014). The head p hosting the post-verbal Theme in its specifier attracts the localiser phrase to create an additional specifier of pP. The Theme \ u N • V K • ‘one tree’ cannot be attracted because it is not in the c-command domain of little S (being in its specifier, not in the complement). Hence, there is no locality violation.

5.3 Locative Inversion without an Implicit Agent

5.3.1 Locative Inversion with Durative/Imperfective -zhe

With -zhe, the LI based on J X j ‘hang’-verbs and [L ‘write’-verbs has a state reading, the LI based on] … X ‘walk’-verbs has a progressive reading, and the LI based on] K z Q J ‘plant’-verbs is ambiguous between a state reading and a progressive reading.

To account for the state reading in (28a), we keep on assuming that the pP is the source of the state meaning. The post-verbal Theme \ u ‘one tree’ N • V K • is introduced by the little p, and the localiser phrase \ X i Q] L O ƒ ‘yard-in’ is merged as the complemen of the null P. Given the lack of event implication in this LI (cf. the ‘again’-test in the section 4.2), we argue that the verb is formed by the root being merged into a popular verbaliser Y (Myler 2016). This verbaliser is stative, and does not introduce an event variable or argument. The result is a semantically beached verb, which behaves like a copula.

---

1 We hypothesise that the localiser phrase fails to get its Case assigned or valued by the little p G X H W K H I D F W W K D W 3 L V Q X O O L Q F R R W H J D W W Z L W K unvalued Case feature on the localiser phrase makes it active for the Probe in T. The movement renders the localiser phrase local to T. This analysis raises the question of why the Case assignment has to care about the overtness of P. Alternatively, the localiser phrase can be introduced by an Applicative head.

(28) (a) \ X i Q] L Ozyhòng -zhe \ t N • shù.
yard-in plant - } Ž ‹ one- | … • tree
Lit. “In the yard is planted one tree.”

(b)

This proposal can provide insights into the following set of data.[2] As shown in (29), the post-verbal Theme arguments are inanimate and only -zhe marking is possible. These inanimate themes are particular, because ‘stone lions’ cannot really ‘sit’, and the ‘moon’ cannot be hung by someone in the sky. The LI with -zhe describes the spatial configuration of these inanimate themes, and there is no real event of ‘sitting’ or ‘hanging’ involved. By contrast, since the LI with -le has event implication due to the eventive verb, the oddness is induced by the events of ‘sitting’ and ‘hanging’ with these inanimate themes.

2 6HH +X ;LDR DQG 0F(QHU\ DQG)HQJ KLV /LX this set of data.

(29) (a) P p Q N X zuò *-le/-zhe \ t G X u V K t V K ¯
doorway sit ‰ ~ ‹ • } Ž œne- | … • stone-lion
“At the doorway sits a pair of stone lions.”

(b) W L — Q V kuà Q J *-le/-zhe \ u O ~ Q míng-yuè.
sky-on hang ‰ ~ ‹ • } Žœne- | … • bright-moon
³ , Q W K H V N \ K D Q J V D E U L J K W P R R Q ´) H Q J

We turn to the progressive reading in the LI based on] … X ‘walk’-verbs and] K z Q J ‘plant’-verbs(30a). As shown in (30b), the post-verbal Theme argument \ u N • V K • ‘one tree’ is merged at the specifier of the dynamic vP. The complex predicate is made of the verb and the location noun \ X i Q] L ‘yard’. Note that the complement of the verb is a locative phrase, and the location noun \ X i Q] L ‘yard’ forms a complex predicate with the verb] K z Q J ‘plant’ and turns the meaning of verb from ‘an event of planting’ to ‘an event of planting the yard’.

(30) (a) \ X i Q] L Oz zhòng -zhe \ t N • shù.
yard-in plant - } Ž ‹ one- | … • tree
Lit. “In the yard is being planted one tree.”

(b)

We argue that the localiser phrase \ X i ‘Qa]L-in O’ is formed at a later stage of derivation, in a separate syntactic Work Space. The reasons for assuming a complex predicate structure like (30b) come from the observation that the location noun can complement] …‘walk’/] K z Q J ‘plant’-verbs elsewhere in the language (cf. non-canonical objects, Barrie and Li 2015, Zhang 2018). As shown in (31), these verbs can take a location noun as a complement. Like the zhe-marking LI based on the same verbs, the occurrence of the durative -zhe in (31) yields a progressive reading.

(31) (a)] X -zhe V K — Q O •
walk - } Ž ‹ montain-road
“be walking the mountain road”

(b) S ý R -zhe F — R F K ý Q J
run - } Ž ‹ playground
“be running the playground”

(c) zhòng -zhe K X — \ X i Q
plant - } Ž ‹ flower-yard
Lit. “be planting the garden”

(d) S Ì -zhe P ý O •
pave - } Ž ‹ horse-road
“be paving the road”

The complex predicate structure proposed in (30b) can give an explanation of why J X j ‘hang’/ [L ‘write’-verbs do not have the LI with a progressive reading. As shown in (32), contrasting with] … X ‘walk’/] K z Q J ‘plant’-verbs, J X j ‘hang’-verbs cannot take location nouns as complement. It follows that J X j ‘hang’-verbs cannot form the structure of (30b). In addition, the reason the LI based on] K z Q J ‘plant’-verbs is ambiguous between a reading and a progressive reading is due to the fact that] K z Q J ‘plant’-verbs can occur in both of the two structures (28b) and (30b) that yield these readings.

(32) (a) *fàng (-zhe) T L i Q J
put - } Ž ‹ wall

(b) [L (-zhe) E Q] L
write - } Ž ‹ notebook

(c) *guà (-zhe) T L i Q J
hang - } Ž ‹ wall

5.3.2 Sing-Verbs

In contrast with] K z Q J ‘plant’-verbs and J X j ‘hang’-verbs, the LI based on F K j Q ‘sing’-verbs is not diagnosed with an implicit Agent (33a). With the perfective -le, the sentence describes a terminated event, not a resultant state. With the durative -zhe, the sentence describes an on-going event of singing.

As shown in (33b), we propose that the verb F K j Q J ‘sing’ takes as complement the post-verbal noun phrase J • M • ‘opera’, which forms a complex predicate with the verb. This nominal predicate enriches the semantics of the verb by turning an event of ‘singing’ into an event of ‘singing the opera’. The localiser phrase is merged in the Spec,vP as an argument. The vP denotes a dynamic event. Above vP, aspectual markers are merged in Asp2 head.

(33) (a) tái-shàng chàng -zhe/-le J • M •
stage-on sing - } Ž ‹/- ‰ ~ ‹ • opera
-zhe: “On the stage is being sung the opera.”
-le: “On the stage was sung the opera.”

(b)

The complex predicate F K j Q J • M • ‘sing-opera’ is predicated of the subject W j L V K j Q J ‘stage-on’. This predicational relation may shed light on the restriction on the verbs. As shown in (34), unlike F K j Q J ‘sing’ and P j L ‘sell’, verbs like W ¯ Q J ‘listen to’ and P • L ‘b

are unable to form LI.[3] We speculate that the complex predicate E K j Q J M • 'sing-opera' in (34a) expresses the property characteristic of the subject W i L V K j Q J 'stage-on', where the complex predicate W J ¯ Q J 'listen to-opera' is not a property of the subject W i L V K j Q J 'stage-on'. In the V H O O E X \ case in (34b), while the complex predicate F K • L P j L expresses the property characteristic of the subject K X] K • Q V K j Q J 'exposition-on', the complex predicate P • L F K • 'buy car' describes the behaviours of the clients. Therefore, verbs that can participate in the complex predicate structure need to describe some properties of the localiser subject.

(34) (a) \ ¯ Q \ X q W ¯ Q J W ¯ Q J
music hall-in sing/listen } Ž ‹ opera
"In the music hall is being sung/*listened to the opera."

(b) K X u] K ý Q V P K j L Q J P ý L
exposition-on sell/buy } Ž ‹
] X u [¯ Q] K u] j R G T H u F K •
most-new make-} ~ car
"In the exposition are being sold/*bought the most recently made cars."

5. Concluding Remarks

We have examined the Locative Inversion based on five subclasses of verbs by studying their occurrence with the perfective -le and the durative -zhe. The interpretive differences and the possibility of an implicit Agent have led us to propose a non-uniform analysis. With the perfective -le, the LI based on 'hang'/'write'/'plant'-verbs has an eventive verb embedding the prepositional pP, yielding the change-of-state meaning. In addition, we argue that the implicit Agent is introduced by Middle Voice (Alexiadou, Anagnostopoulou and Schäfer 2015), which should be distinguished from the Mandarin canonical passive with bèi. With the durative -zhe, the LI based on 'hang'/'write'/'plant'-verbs denotes state. We have argued that the root of the verb is adjoined to the light copular verbaliser v. As a result, the verb behaves like a copula, and takes the state-denoting pP as complement. With the durative -zhe, the LI based on 'walk'/'plant'-verbs has a dynamic vP, which takes the location noun as complement and the post-verbal theme as argument. The LI based on 'sing'-verbs has an unaccusative structure. The result shows that the Mandarin Locative Inversion can take place when the language has several strategies to produce structures without an overt Agent.

---

3 There is a contrast between sell and E X \ in the formation of English middle constructions (Fagan 1992).

(i) a. The new Saramago sells like water in a desert.
b. * The new Saramago buys Z L W K J U H D W G L v F X O W \ W K H G L V W U L

Works Cited

Alexiadou, Artemis, Elena Anagnostopoulou and Florian Schäfer. 2015. ([WHUQDO $UJ
LQ 7UDQVLWLYLW\ $OWHUQDWLRQV $ /D\HUHGQHVV $SSU
Alexiadou, Artemis, and Edit Doron. 2012. “The Syntactic Construction of Two
Non-Active Voices: Passive and Middle.” 1– R XUQDO RI /LQJXLVWLFV
Barrie, Michael, and Yen-Hui Audrey Li. 2015. “Analysis vs. Synthesis: Objects & Ki
6\QWD[LQ D &URVV /LQJXLVWLF 3HUVSHFWLYH, edited
Simpson, and Wei-Tien Dylan Tsai, 179–206. Oxford: Oxford University Press.
Bobaljik, Jonathan David. 1995. “ 0 R U S K R V \ Q W D [7 K H 6 \ Q W D [R I
Ph.D. diss., MIT.
Chomsky, Noam. 1995 7 K H 0 L Q L P D O L V W 3 U R J U D P. Cambridge, MA: MI
Chomsky, Noam. 2000. “Minimalist Inquiries: The Framework.” 6W H S E \ 6 W H S (V V
R Q 0 L Q L P D O L V W 6 \ Q W D [L Q + R Q R U R I + R Z D U G / D V Q L
and J. Uriagereka, 89–155. Cambridge: MIT Press.
Chomsky, Noam. 2001. “Derivation by Phase.” In . H Q + D O H $ / L I H L Q / D Q J
by M. Kenstowicz, 1–52. Cambridge, MA: The MIT Press.
Fan, Fanglian. 1963. “Existentials.” 6W X G L H V R I & K L Q H V H / D Q J X D J H
Gu, Yang. 1992. “On the Locative Existential Construction in Chinese.” In 3 U R F H H G L Q
R I W K H W K : H V W & R D V W & R Q I H U H Q F H R Q) R U P D O
í 6 W D Q I R U G & $ 7 K H 6 W D Q I R U G / L Q J X L V W L F V
Hoekstra, Teun, and René Mulder. 1990. “Unergatives as Copular Verbs; Locational and
Existential Predication.” / L Q J X L V W L 7 – 79. Y L H Z
Hu, Wenze. 1995. “Verbal Semantics of Presentative Sentences.” < X \ D Q < D Q M
100–112.
Huang, Cheng-Teh. James. 1987. “Existential Sentences in Chinese and (In-)definite
ness.” In 7 K H 5 H S U H V H Q W D W L R Q Edited by Eric Reuland and Alice
ter Meulen, 226–253. Cambridge, MA: MIT Press.
Huang, Cheng-Teh. James., Yen-Hui Audrey Li and Yafei Li. 2009. 7 K H 6 \ Q W D [
& K L Q H V H Cambridge: Cambridge University Press.
Kratzer, Angelika. 1996. “Severing the external argument from its verb.” In 3 K U D V H 6 W U X
D Q G W K H / H [L F R Q H G L W H G E \ - 5 R R U \ F N D Q G / = D U
Legate, Julie Anne. 2014. 9 R L F H D Q G Y / H V . MIT Press. U R P $ F H K Q H V
Levin, Beth, and Malka Rappaport Hovav. 1995. 8 Q D F F X V D W L Y L W \ $ W W K
6 H P D Q W L F . Cambridge, MA: MIT Press.
Levinson, Lisa. 2011. “Possessive WITH in Germanic: HAVE and the Role of P.”
6 \ Q W D [í
Li, Yen-hui Audrey. 1990. 2 U G H U D Q G & R Q V W L W X H Q F \ L Q 0 D Q G D
Kluwer Academic Publishers.
Lin, Jo-Wang. 2002. “Aspectual Selection and Temporal Reference of the Chinese Aspec
tual Marker Zhe.” 7 K H 7 V L Q J + X D - R X U Q D O R I & K L Q H V H 6 W

Lin, Tzong-Hong. Jonah. 2008. “Locative Subject in Mandarin Chinese.” 1 D Q] D Q / L Q J X L V W L F V 4: 69–88.
Liu, Feng-hsi. 2007. “Auxiliary selection in Chinese.” In 6 S O L W $ X [L O L D U \ $ & U R V V / L Q J X L V W L H G L 3W HU S SH F5 D L- YO H $ U D Q R Y L F K i John Benjamins.
Myler, Neil. 2016. % X L O G L Q J D Q G , Q W H U S U H W L Q J 3 R V V H V V L MIT Press.
Pan, Haihua. 1996. “Imperfective Aspect zhe, Agent Deletion, and Locative Inversion in Mandarin Chinese.” 1 D W X U D O / D Q J X D J H D Q G / L Q J X L V W L F
Paul, Waltraud, Yaqiao Lu, and Thomas Hun-Tak Lee. 2019. “Existential and Locative Constructions in Mandarin Chinese.” / L Q J X L V W L F 5 H Y L HíZ
Smith, Carlota S. 1997 7 K H 3 D U D P H W Dordrecht: Kluwer.
Svenonius, Peter. 2010. “Spatial P in English.” In 0 D S S L Q J (The Oxford Handbook of 3 3 V of Syntactic Structures 6), edited by Guglielmo Cinque and Luigi Rizzi, 127–160. New York: Oxford University Press.
Travis, Lisa. 2010. Q Q H U $ V S H F W 7 K H $ U W L F X O D W L R Q R I 9 3
Tsai, Wei-Tien Dylan. 2008. “Tense Anchoring in Chinese.” Q J X D 118: 675–686.
von Stechow, Arnim. 1996. “The Different Readings of Wieder ‘Again’: A Structural Account.” - R X U Q D O R 13(2): 87–138. doi: 10.1093/jos/13.2.87.
Williams, Adina. 2016. Modificational and Predicate Adjunct Uses of the Mandarin Locative Preposition zài. Paper presented at ConSOLE XXIV, January 6–8, 2016.
Xiao, Ricard, and McEnery, Tony. 2004. $ V S H F W L Q 0 D Q G D U L Q & K L Q H V H 6 W X G \. Amsterdam and Philadelphia: John Benjamins.
Yang, Suying. 1999. “The Unaccusativity Phenomenon: A Study on the Relationships between Semantics and Syntax.” D Q J G D L \ X 1 3 Q 4 3 X H
Yang, Suying, and Pan, Haihua. 2001. “A Constructional Analysis of the Existential Structure.” In 6 W X G L H V L Q & K L Q H V H / L Q J X L V W L F V edited by H-H. Pan, 189–208. Hong Kong: Linguistic Society of Hong Kong.
Yeh, Meng. 1993. “The Stative Situation and the Imperfective zhe in Mandarin.” - R X U Q D O R I W K H & K L Q H V H / D Q J X D J H 7 H D F K H U V $ V V R F L D W L R
Yu, Ning. 1995. “Towards a Definition of Unaccusative Verbs in Chinese.” In 3 U R F H H G L Q J V R I W K H W K 1 R U W K $ P H U L F D Q & R Q I H U H Q F H J. Camacho and L. Choueiri, 339–353. Los Angeles: Univerity of South California.
Yuan, Boping. 1999. “Acquiring the Unaccusative/Unergative Distinction in a Second Language: Evidence from English-Speaking Learners of L2 Chinese.” / L Q J X L V W 37: 275–296.
Zhang, Niina Ning. 2018. “Non-Canonical Objects as Event Kind-Classifying Elements.” 1 D W X U D O / D Q J X D J H 36(4): 1395–1437. F 7 K H R U \
Zhang, Nina Ning. 2019. “Appearance and Existence in Mandarin Chinese.” 6 W X G L H V L Q & K L Q H V H / 40(2): 101–140. F V

Case Changing and Case Maintaining Movements in Dependent Case Theory: Dative Extraction in Hungarian

Mark Newson [a] and Krisztina Szécsényi [b]

[a, b] Eötvös Loránd University, Budapest, Hungary

[a] newson.mark@btk.elte.hu, [b] szecsenyi.krisztina@btk.elte.hu

Abstract: In Hungarian, when a possessor moves from its position within the np to a higher position in the DP, it changes its case from nominative to dative. But when it moves from this higher position, it maintains this dative case. This paper seeks to account for why some movements allow case change while others do not. Our account is set within Baker's (2015) Dependent Case Theory, though with modifications which enable us to achieve our goal. We claim that the Invisibility Principle (case assignment makes a DP invisible to the case system) is central to the theory and use it to justify the distinction between Baker's hard and soft domains. A soft domain is defined as one for which its unmarked case is not assigned until a hard domain is evaluated. Thus, DPs in soft domains earmarked for unmarked cases may get a different case if they move.

Keywords: Dependent Case Theory; unmarked case; possessor extraction; hard and soft domains

1. Introduction

The aim of the paper is to provide an explanatory account of the following cross-linguistic observation: occasionally the case of a DP changes after movement. Since existing accounts fail to reliably predict when case change takes place and when it does not, we offer an alternative that argues for a modification of Baker's (2015) Dependent Case Theory.

Our analysis is based on the following empirical facts from Hungarian:

I. Movement of a nominative possessor within a DP changes the case of the moved possessor into dative without any change in meaning:

(1) (a) a ¿ ~ kalap-ja
the boy. ‡ ˆ † hat- Œ € ‰ ˆ Œ Œ
‘the boy’s hat’

(b) a ¿ ~ Q D N kalap-ja
the boy- } z • the hat- Œ € ‰ ˆ Œ Œ
‘the boy’s hat’

In the well-known examples in (1) (Szabolcsi 1994, 1984), the position of the definite article is a reliable indicator of the position of the possessor: in sentence (1a) the possessor is behind the article, in (1b) it is before it, suggesting a higher external position. Correspondingly, in (1a) the possessor must be nominative, while in (1b) it is obligatorily in dative case.

II. Movement of a dative possessor/subject from DPs (2) and TPs (3) leads to no change of the case of the moved DP.

(2) > $ ¿ ~ ¡Q/C*[D N$@¿ ~ @ elveszett [t_i a N D O D S M D @
the boy- } z • the boy. ‡ ˆ † got.lost the hat- Œ € ‰ ˆ Œ Œ
‘The boy’s hat got lost.’

(2) shows is that it is possible for the possessor to be extracted in Hungarian, but only if the possessor is dative.

(3) (Péter-nek) nem kell (Péter-nek) el-men-ni-e.
Peter- } z • not have.to away-go- , ‡ • Œ €
‘Peter does not have to leave.’

(3) shows that the subject of an inflected infinitive is dative clause internally (cf. Tóth 2000) and can move to the left periphery of the selecting verb. The dative case form remains unchanged.

Having presented the empirical focus of our paper, in section 2 we argue in support of the movement analysis and against a base account in terms of base generation. Section 3 presents background information on Baker’s (2015) Dependent Case Theory (DCT) highlighting the notion of soft and hard domains and discusses the results of earlier work on a DCT account of Hungarian (Newson and Szécsényi 2020). Section 4 looks more closely at Baker’s (2015) stance on the relationship between case and movement and identifies a number of problems in his analysis. In section 5 we present an alternative analysis, which makes the soft vs. hard domain distinction meaningful and accounts for the distribution of case change.

2. Case and Movement in Hungarian

One instance of our claim that movement from a dative position does not change the case of the DP concerns DP external possessors. In arguing that we are dealing with movement (as opposed to base generation) in the relevant structures presented in section (1) we rely on den Dikken (1999) and É. Kiss (2014). These use evidence from the anti-agreement facts of the possessive DP paradigm to distinguish between two possibilities for external possessors. The relevant observation is that in possessive DPs with a third person plural possessor, the possessor and the possessee cannot agree as they do in other cases (see, for example, (1)). When the possessor is pronominal, while the meaning is plural, it can only appear in the singular form , although the possessee carries the plural agreement (4a). When the possessor is a nominative lexical DP, it appears in its plural form, and the agreement marker on the possessee has to be third person singular (4b):

(4) (a) az N kalap-juk
the Œ € ‰at-. ‰ … ‰ ˆ Œ Œ
'their hat'

(b) a ¿ ~ N kalap-ja/*kalap-juk
the boys at- Œ €/hat- ‰ …
'the boys' hat'

Interestingly, when a plural lexical possessor is extracted (and hence realized as a dativeDP), the possessee can bear either agreement form (5). In order to account for this, den Dikken (1999), and, following the proposal therein, É. Kiss (2014) argue for different derivations. The pattern showing anti-agreement can only arise if the possessor moves from within the possessive DP (5a). In order to explain (5b), where the possessee has a form it would have if it (anti-)agreed with a pronoun, it is assumed that this is indeed the case: within the possessive DP there is a pro with the lexical possessor being base generated as an affected argument of the selecting verb[1] (5b)[2]

1 Here, the dative case of the base generated external possessor cannot come from within the possessive DP. We assume that is has another source entirely and is an inherent case assigned by the applicative head which introduces it.

2 A reviewer points out that plural possessive DPs can be interpreted as having either collec W L Y H S R V V H V V R U V R U G L V W U L E X W H G R Q H V G L Q u H U D H Q W R U K D V for each boy or the boys collectively owned the hat. Moreover, according to their intuition, there is a preference for the distributed reading in the case of (5a). We have nothing to add to these comments at this point.

(5) (a) $ ¿ ~ N$_i$ Q D elveszett [$_{DP}$ a [$_{NP}$ t$_i$ kalap-ja @ @
the boys-}z• got.lost the hat- Œ € ‰ ˆ Œ Œ
‘The boys’ hat got lost.’

(b) $ ¿ ~ N$_i$ Q D elveszett [$_{DP}$ a [$_{NP}$ pro$_i$ kalap-juk @ @
the boys-}z• got.lost the hat-3 ‰ … ‰ ˆ Œ Œ
‘The boys’ hat got lost.’

That is, anti-agreement can be used as a diagnostic for movement. It is the data showing anti-agreement like the one in (5a) that are relevant for the purposes of the present paper. The pattern in (5b) will not be discussed further.

We now take a closer look at whether familiar dichotomies in case and movement can be used to account for the distribution of case-changing and non-case-changing move ments. One of the obvious candidates is the distinction between inherent and structural case, as inherent case is known to be maintained under movement and structural case can be acquired by a DP in its landing site. Though Hungarian does have inherent datives, assuming a difference in case type will not explain case-change as the constructions in question involve structural datives only, as indicated by the fact that the relevant DPs are not restricted to a specific interpretation. Furthermore, as argued in Newson and Szécsényi (2020), the fact that the dative subjects of inflected infinitives show clear parallels with the nominative subjects of finite clauses, both cases emerging on subjects in transitive and intransitive contexts ((6) and (7)), demonstrates that these cases are not only structural but specifically unmarked.

(6) (a) Muszáj [Péter-nek K D] D P H Q Q L H @
must Peter-}z• home- € ˆ , ‡ • Œ €
‘Peter must go home.’

(b) Tud-om [Péter K D] D P H Q W @
know-1 Œ € Peter. ‡ ˆ † home- € ˆ ‰ Œ •
‘I know that Peter went home.’

(7) (a) Nem szabad [Péter-nek meg-néz-ni-e ez-t a ¿ O P H W @
not allowed Peter-}z• ‰ watch- , ‡ • this- z | | the ¿ O P z | |
µ 3 H W H U L V Q R W D O O R Z H G W R Z D W F K W K L V ¿ O P

(b) Tudom [hogy Péter meg-néz-t-e ez-t a ¿ O P H W @
know-1 Œ € that Peter. ‡ ˆ † ‰ watch- ‰ this- z | | the ¿ O P z | |
µ , N Q R Z W K D W 3 H W H U Z D W F K H G W K L V ¿ O P ¶

We have already seen that in possessive structures the case form of the possessor depends on its position: lower possessors are nominative, higher ones, dative. Both cases are necessarily structural. Still, it is only nominative case that changes; dative does not. In conclusion, the inherent vs. structural case distinction cannot account for when case changes after movement.

Another potential candidate for predicting when case change takes place and when it does not is the A- vs. A-bar movement dichotomy. Under traditional assumptions, A-movement involves moving to a position to which case is assigned while A-bar movement does not. So, a DP undergoing raising or passivisation, for example, will receive a case associated with its landing site while a DP undergoing wh-movement will maintain the case it was assigned before it moved to spec CP. However, much current work questions the legitimacy of this distinction on empirical and conceptual grounds. Conceptually the distinction between argument and non-argument positions is hard to maintain under current assumptions. There are movements which do not cleanly fall into one or the other type: scrambling, for example, resists such an account in terms of this distinction (Webelhuth 1992). Specifically concerning the Hungarian data, there is little reason to assume an A/A-bar distinction between the movement from the nominative and dative positions: both are movements from positions to which case is assigned, though both fail to qualify as A-bar movement as none of the features associated with this can be convincingly argued to trigger either movement. If anything, this movement could be identified as one which makes further movement out of the DP possible.

3. Background

We have seen in the previous section that a movement account of the constructions in question is feasible, but at the same time case change/maintenance does not correlate with either case or movement type. Clearly, there must be something else that accounts for its distribution, but from the standard perspective it is far from obvious what our next candidate should be. The main claim of this paper is that an alternative view of case, Baker's (2015) Dependent Case Theory, fares better, provided that certain modifications are introduced.

In order to proceed, we need to introduce some background information on the core notions of Dependent Case Theory in general, and, more specifically, previous assumptions on the case domains of Hungarian in a DCT framework.

3.1. Dependent Case Theory

One of the core components of DCT is the case hierarchy of Marantz (1991), according to which DPs are assigned case in an order that depends on case type: lexical case, case with a consistent semantic value is assigned first. After this comes dependent case, which is assigned to one of two case competitors in a local domain. Unmarked case is assigned

to those DPs that have no lexical case or dependent case assigned to them.[3] Importantly, unmarked case is not necessarily morphologically unmarked, and can vary even in the different domains of one and the same language.

In Marantz's system, it was assumed that after a DP is assigned case, it plays no further role in the system; either to receive another case or to provide the conditions for dependent case to be assigned to another DP. Baker, however, marginalizes this idea by making it dependent on a parameter setting. For languages like Icelandic, in which quirky subjects trigger nominative case assignment to objects, the parameter is set so that once the subject is assigned inherent case, it drops out of the competition and the object is then eligible for unmarked case, being the sole remaining contender in the domain. However, for some languages, such as Warlpiri and Burushaski, some predicates which have an inherent dative object also have an ergative subject, suggesting that for these languages the inherent case marked object is still visible to enable the high dependent case to be assigned to the subject.

We will discuss problems for Baker's proposal in section 4.2, but for now we will simply reject it and instead propose to elevate Marantz's original proposal the status of a core principle of the framework. Let us call it the Invisibility Principle:

(8) The Invisibility Principle
Once DPs are assigned case, they play no further role in the case system

Another key component of the theory is the notion of case domains. These have two important functions. First, they limit the DPs competing for assigned case; only those which appear in the same domain can determine the conditions of dependent and unmarked case assignment. Thus, while the presence or absence of an object can determine which case is assigned to the subject of the object's clause, it will never affect the case of a subject in a different clause. Second, dependent and unmarked cases are specified for specific domains and can vary both cross-linguistically and intra-linguistically, e.g. we often find different cases in nominal vs. verbal projections.

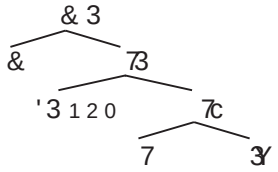
Baker (2015) equates case domains with the spell-out domains of phase theory (Chomsky 2001). We will be in a better position to understand the problems this leads to after a discussion of how unmarked nominative and dative domains work in Hungarian.

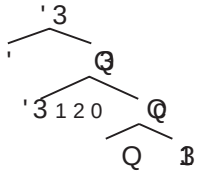
3.2. Unmarked Nominative and Dative in Hungarian

This part of the paper builds heavily on the conclusions reached in Newson and Szécsényi (2020), the aim of which is giving a consistent account of the distribution of nominative and structural dative DPs in Hungarian in a DCT framework.

[3] 'LuHUHQW FDVH V\VWHPV HPHUJH GHSHQGLQJ RQ ZKHWK assigned the dependent case.

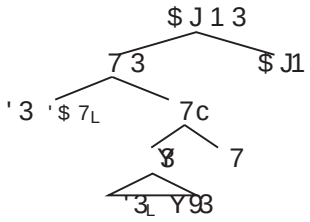
The structure we assume for clauses in (9) is standard. For the possessive DP (10), we assume that the possessor is introduced by n, similar to how v introduces the agent.

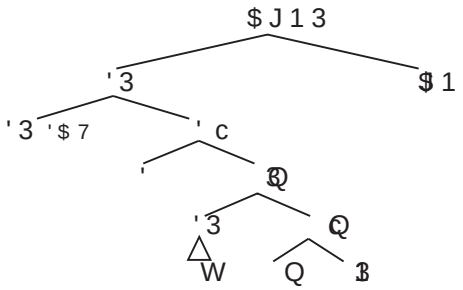




The complements of C and D are therefore domains with unmarked nominative case. CP and DP are the typical extended verbal and nominal projections, in the sense of Grimshaw (2005), and therefore the unmarked nominative domains of Hungarian are the complements of the heads of canonical extended projections. In these domains, the subject/possessor is the highest, if not only, DP and therefore is eligible for unmarked case assignment.

The dative DPs we are concerned with in the present paper are all associated with unmarked dative domains in Newson and Szécsényi (2020). Among others, we syste matically find unmarked dative case in the complements AgN, the non-finite agreement head that appears in possessive DPs and inflected infinitives and differs from the finite agreement paradigm.





In the inflected infinitive in (11) (see (6a) for example), the subject sits in a similar position to that of other clauses, the only difference being the presence of AgN. This defines TP as an unmarked dative domain and hence the subject, being eligible for unmarked case, gets dative. In the possessive DP in (12) (see (1b)), the possessor has moved from its position in (10) to the specifier of the DP complement of AgN, where it receives dative for the same reason as the subject of the inflected infinitive.

3.3. Hard and Soft Domains

In the previous section we saw how the DPs occupying the specifier position of complements of heads determining unmarked nominative and dative domains end up in their respective case forms. Now we can move on to the problems that emerge for Baker's (2015) approach to case domains.

To recap, for Baker case domains are the spell-out domains of phase theory. This often leads to the right predictions, but, unmoderated it runs into a systematic and rather disturbing problem: VP should be a case domain, as v is a phase head universally. This is problematic because in most standard case systems the specifier of VP, when present, interacts with DPs in the clausal domain to determine dependent case assignment: accusative to the object or ergative to the subject. It therefore seems that movement and

4 A reviewer points out that dative possessors can also appear postnominally:

i. a két szép régi kocsi-ja a nagyi-nak
the two nice old car- Œ ‰ the Œ grandmother- } z •
"the grandmother's two nice old cars"

Given that dative case, once assigned to a possessor does not change on movement, a straightforward account of postnominal datives would be to employ a movement similar to extraposition. Similar phenomena can be found in certain Hungarian PPs where a dative nominal not only can, but must be extraposed.

ii. Mellett-e áll-t-am Péternek. (vs. Péter mellett áll-t-am.)
beside- Œ € ‰ ^ Œ Œ stand- ‰ z Œ • Peter Œ € ‡ Peter } z • stand- ‰ z Œ • Œ €
Both: "I was standing next to Peter."

Space limitations prevent us from going into these observations further in this paper.

case assignment part ways at this point: whereas for movement we have ample evidence for vP being a phase, for case assignment we need a larger local domain, at least in some cases. In other cases, however, as Baker points out, assuming the VP to be a case domain is advantageous. He argues that in some languages we can identify cases, both unmarked and dependent, which are specific to the VP. For example, he claims that Sakha has a high dependent dative and Finnish has an unmarked partitive in the VP. As it is one of the functions of domains to distinguish the different structural cases assigned to DPs in different positions, this is a strong reason to think of VP as a domain in these cases. Therefore, it seems that languages differ in whether VP counts as a domain or not.

To account for this cross-linguistic variation, Baker introduces a distinction between soft and hard case domains. The two differ in their effects on case. Hard domains behave as expected with their contents being spelled out on the merger of the phase head. Therefore, contained DPs have fixed cases and they do not participate in determining the case assignment to DPs in subsequent domains. In soft domains, contained DPs remain active after spell-out, and therefore are available to receive cases assigned at a higher domain and can act to determine dependent case assignment at that higher domain. Thus, if the VP is soft, the object can be assigned accusative at the TP level or determine ergative case assignment to the subject.

While this distinction is necessary for Baker, it remains descriptive, without explanatory content and does not naturally fall out from either case or phase theory. It is an aim of the present paper, in elevating the status of the Visibility Principle, to place soft domains as a proper component of DCT with explanatory power, resulting in an account that is also more in line with minimalist assumptions.

7. Baker on Case and Movement

Baker sees the relationship between case assignment and movement as a timing issue: which precedes and which follows? Originally standard case theory saw the Case Filter operating at S-structure, allowed the possibility of moving caseless DPs into case positions. Other movements seemed to carry an already assigned case to positions not associated with case assignment. Baker maintains that this distinction is still relevant within Dependent Case Theory, though not necessarily applying to the same phenomena which motivated it under standard assumptions. The question is that if case domains are equated with Spell-Out domains, and case assignment is part of spelling out a DP, how can it be delayed until after movement?

Baker proposes several different mechanisms to deal with various instances related to this issue. For instance, he proposes that adjuncts are spelled out in a second round within each phase, following the spell out of all non-adjuncts. He uses this to account for why both the subject and the extracted possessor in Japanese end up with marked nominative:

(13) John-ga [– R W R R V D Q s@n-da D
John- † ‡ ˆ † father- † ‡ ˆ † die- ‰ z Œ •
'It is John whose father died.'

On the assumption that the extracted possessor is adjoined in a subject position, Baker argues that the subject in (13) is assigned marked nominative case because, at the point it is evaluated, the adjoined possessor is not considered. As there is no DP c-commanding the subject, it is eligible for marked nominative.[5] Subsequently, the extracted possessor is considered and as there is no c-commanding DP for this it also gets marked nominative.

For our paper, however, the most important assumption Baker makes concerns how different cases interact with movement. His claim is that dependent case, once assigned, cannot be altered if a DP is evaluated in another domain. But unmarked case may be overwritten in such circumstances. A DP can be evaluated in more than one domain under two conditions: i) when it is inside of a domain, and so it remains visible after the domain is spelled out, and ii) when it moves from one domain to another. Thus, dependent case will be maintained, and unmarked case may be changed, under movement.

Evidence that dependent case is maintained under movement comes from raising of ergative subjects (14) and moving internal arguments with structural dative from VP to subject position (15):

(14) (a) ‘E lava [‘o ako ‘e Pita ‘a e lea I D N D 7 R Q J D @
z Ž ‘possible | ˆ † ‰ learn ~ ‹ € Peter z { Œ the language Tongan
‘Peter can learn Tongan.’

(b) ‘E lava ‘e Pita [‘o ako – ‘a e lea I D N D 7 R Q J D @
z Ž ‘possible ~ ‹ € Pita | ˆ † ‰ learn z { Œ the language Tongan
‘Peter can learn Tongan.’

(15) Ucuutal-ga student-na tiij-bet-ter.
teacher-} z• student- ‰ … V X v F H ‡ ~ € z ˆ ‹ ‰ Œ
‘The teacher doesn’t have enough students.’

In (14) we see a Tongan example of subject raising where we observe that the raised subject maintains its ergative case and (15) illustrates a Sakha diadic unaccusative verb in which the dative DP is moved to subject position maintaining its dative case.[6]

5 Baker analyses marked nominative languages as involving the following assignment conditions:

i. Assign a DP marked nominative if it has no c-commanding competitor.

6 Baker maintains that the higher argument in the VP in Sakha receives structural dative, not inherent case, as the dative marked argument is not restricted to a particular semantic interpreta W L R Q D Q G D S S H D U V L Q D Q X P E H U R I G L u H U H Q W F R Q V W U X F

As for unmarked case being overwritten with movement of the DP to another domain, Baker provides two examples. The first involves his account of differential object marking in Sakha. In this language, definite DPs receive accusative case while indefinite DPs are unmarked. The data below indicate that the definite DP is raised out of the VP, and hence precedes the VP adverb, while the indefinite DP remains inside the VP:

(16) (a) Masha [VP türgennik salamaat V L H W H @
Masha quickly porridge eat-‰ z Œ • Œ Œ
'Masha ate porridge quickly.'

(b) Masha salamaat-y türgennik [VP V L H W H @
Masha porridge- z | | quickly eat-‰ z Œ • Œ Œ
'Masha ate the porridge quickly.'

On the assumption that the VP is a hard domain in Sakha, Baker accounts for the unmarked case on the VP internal indefinite object straightforwardly: when the VP is spelled out, the only DP in the domain is the object and hence it is assigned unmarked case. Remaining inside the VP, this object plays no further role in the case system and hence the unmarked case is fixed. However, the definite object moves out of the VP into the higher TP domain. The unmarked case that it was assigned in its VP internal position is then overwritten by the dependent (accusative) case as in the higher domain the object is in a position to interact with the subject.

The second instance of unmarked case being overwritten after movement concerns the Japanese data given in (13). First Baker argues that, in some languages, genitive case is the unmarked case of the nominal domain. The mark of an unmarked genitive is the possibility of having two genitive DPs within a single possessive DP. As the following shows, Japanese can be taken to have an unmarked genitive:

(17) yuubokumin-no toshi-ɲo hakai
nomad- € ~ ‡ city- € ~ ‡ destruction
'the nomad's destruction of the city'

As we saw in (13), when a possessor is extracted from the possessive DP it loses its genitive case and is overwritten with the case determined for it in the higher domain. The extraction of a possessor with dependent genitive case, however, would not allow this case change to take place. Baker cites Cuzco Quechuan as an example of this.

4.1 Empirical Problems for Baker's Proposal

Although Baker discusses several languages which, by his diagnostic, have unmarked genitive and a number of languages which allow extraction from the possessive DP

accompanied by case change on the extracted possessor, the only example he discusses in which these properties coincide is Japanese.[7]

However, one might wonder if this is enough data on which to form a generalization. In fact, one can find examples which appear to go against Baker's generalization. For some speakers of Greek, it is possible to have two genitives in a single DP, indicating that, for them, genitive is unmarked:

(18) I the-‡^† I R W R J U D ¿ D photograph-‡^† tis the-€~‡ Maria-€~‡ Marias tu the-€~‡ John-€~‡ Yann¡ en is N D O L W € H U L good

pu than tu the-€~‡ Kosta Kosta-€~‡

'John's photograph of Maria is better than the one of Kosta.' Natalia Pavlou p.c.

Moreover, Greek also allows possessor extraction, though without changing the genitive of the possessor:[8]

(19) (a) Tinos whose-€~‡ mu my ipes said-‡} pos that dhiavases read-‡} to the-‡^† vivilo. book-‡^†

'Whose book did you tell me that you read.'

(b) Tis the-€~‡ MARIAS Maria-€~‡ thelo want-1s•. to forema. the-‡^† dress-‡^†

'I want Maria's dress (not Eleni's)' Ntelitheos (2002)

The Hungarian data discussed at the beginning of this paper constitute a problem. As was demonstrated earlier, a possessor moving from the lower DP internal position to the higher one loses its unmarked nominative case, following Baker's prediction. However, when the possessor is extracted from the higher position it maintains its dative case. As dative is the unmarked case of the higher nominal domain, we would expect it to be visible for further case assignment in domain it is moved into. Yet, as we have seen, a moved dative DP never changes case.

Another empirical problem facing Baker's approach is that universally some move ments never involve case change regardless of whether the affected DP has dependent

7 Perhaps not surprisingly, Korean is a language demonstrating similar facts to Japanese. Korean allows double genitives (An, 2014) and extracted possessors are assigned nominative rather than maintaining their genitive case (Nakamura, 2002).

8 It remains to be discovered whether all those people who accept the double genitive construction also allow possessor extraction without case change.

or unmarked case. For example, DP movement out of the TP into the left periphery of the clause necessarily extracts the DP from a case domain, as TP is a spell out domain universally. A left peripheral DP therefore could in principle be evaluated in the higher domain and, if its original case were unmarked, could be assigned another case. But to our knowledge, movements traditionally associated with spec CP, such as wh-movement or first position topics in V2 languages, never change the case of the extracted DP, regardless of what its original case is. This is an important observation as, given that TP is not only a spell out domain, but also universally a hard domain, we claim that it is not only the type of case a moved DP bears which determines whether the case may be changed, but the type of domain that it moves out of. Specifically, only a DP with unmarked case moving out of a soft domain can change its case. All DPs, whether assigned dependent or unmarked case, moving out of a hard domain, maintain their original cases.

4.2. Conceptual Problems

Besides the empirical problems with Baker's account of case changing movements, there are also a number of conceptual problems that Baker's theory faces. One of these concerns the distinction between hard and soft domains. As we pointed out earlier, this distinction is necessitated by the identification of case domains as spell out domains. However, there is nothing in Dependent Case Theory which would predict either the distinction itself or the particular properties of soft domains. Why should some domains be such that already spelled out DPs within them remain active in higher domains?

Moreover, the necessity to introduce the difference between soft and hard domains was introduced by unifying case and spell out domains. Before this, the distinction was unnecessary as no observations concerning movement motivated it. It therefore seems that the distinction relates solely to case phenomena. Unless we can identify something specifically related to case from which it can be made to follow, this discrepancy is clearly problematic for the assumption that the processes of movement and case assignment are restricted by the same structural conditions.

A second somewhat related problem concerns Baker's claim that DPs with unmarked case can be reassigned another case at some later point. As we mentioned above, Baker claims that what we have referred to as the Invisibility Principle is parameterized across languages. Presumably, then, those languages which allow a further case to be assigned to an already case marked DP should have the parameter set to the value that allows case marked DPs to remain visible. But if this is the way to deal with these observations, one would expect consistency within languages and all case marked DPs should either be visible or invisible in subsequent domains. Baker does not attempt to verify this, but it is easy to show that in fact it does not hold. For example, Japanese extracted possessors must remain visible after they have been assigned unmarked genitive so that they can be reassigned marked nominative. Thus, Japanese sets the parameter to the value that allows case marked DPs to continue to take part in the case system. However, Japanese

also has dative quirky subjects, and these cannot trigger accusative case on an object. Instead, the object must appear in nominative, indicating that the inherent case marked DP is invisible:

(20) John-ni sore-ga mie-ta Kishimoto (2013)
John-}z• that-‡ˆ† see-‰z Œ•
"John saw that".

It seems therefore that Japanese adopts the parameter setting which prevents case marked DPs playing a further role in the system, in contradiction to the conclusion just reached.[9]

The Hungarian data is perhaps even more problematic for Baker's parameterized account. Here within a single language movement of the same DP produces different effects, depending on the movement: from the lower DP internal domain to the upper one there is case change and from the upper domain to DP external positions there is no case change. Clearly this could not be accounted for in terms of the kind of macro parameter that Baker suggests.[10]

10. Analysis

We want to address both the empirical and conceptual problems identified in the previous section and see a unified way of doing so. Rather than relegating the Invisibility Principle to the parameterised margins, as Baker does, we propose to elevate it to a central position in the theory, using it not only to conceptualise the distinction between soft and hard domains but to account for why some movements change the case of the moved DP while others maintain it.

Our proposal is rather simple:

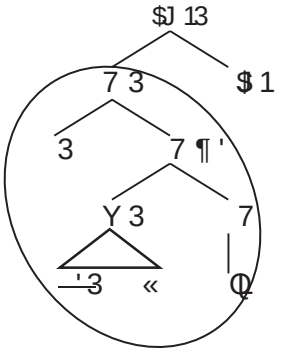
(21) A soft domain is one in which unmarked case is not assigned.

9 Baker could claim that Japanese dative subjects are PPs and therefore do not count as competitors. This would seem unlikely, however, as Kishimoto demonstrates that dative subjects behave like nominative subjects and unlike obliques, suggesting a DP status for them.

10 It remains to account for the Warpiri and Burushaski data in which dative objects do not prevent the subject from receiving dependent ergative case. The fact that in both languages this phenomenon is restricted to certain verbs, indicates that ergative might be lexically assigned in these cases. Baker rejects this as these ergative subjects do not have a uniform argument role. +RZHYHU :RROIRUG SRGLQWVHUHQWHWEKHDWZHWKHQHULEKLVUEHQ WR VSHFL¿F DUJXPHQWV DLQGGL RVH\[QIFDDWLDFVSHUZRKSLHFUKWL\VRD QF it may be that ergative is lexical for the verbs in question.

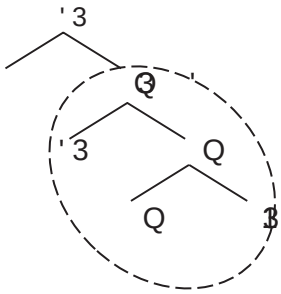
By this, we do not intend to claim that soft domains do not have an unmarked case associated with them. Clearly, they do. What we intend is that the assignment of the unmarked case associated with a soft domain is delayed until the spell out of the next hard domain. This will leave some DPs in some domains unvalued for case feature and therefore, by the Invisibility Principle, still visible in further structural processing. At the level of the hard domain, depending on the structural conditions holding there, visible DPs from the spelled out soft domain can either be assigned the unmarked case defined for that domain or another case associated with the hard domain.

To demonstrate how this works, let us take some examples from the Hungarian data we have been considering. We start with the dative subject of the inflected infinitive. Our story starts at the point when the non-finite agreement is merged with its TP complement in which the subject has already undergone a movement from its vP internal position. According to the analysis given in Newson and Szécsényi (2020), AgN has a soft phase head and therefore TP is a case domain. Here we add the claim that AgN has a hard phase head and therefore TP is a hard case domain, as indicated by the solid circle around it in the following diagram:

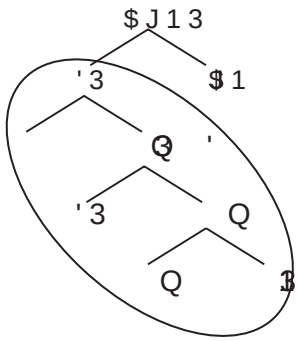


At this point all visible DPs will be assigned case and given that the subject in its TP position has no c-commanding possessor it is determined to receive unmarked case. Moreover, as it sits in the TP domain, the case it will receive will be the one defined for that. As previously argued, the unmarked case for the complement of AgN is dative. Henceforth, as this DP has been assigned case, it can play no further role in the case system by the Invisibility Principle. This means that it cannot itself be assigned another case and hence the extracted dative subject will remain dative.

Compare this to what happens to a DP extracted from the lower possessor position in the DP. We start at the point where the D is merged into the structure, taking the possessive nP as its complement:

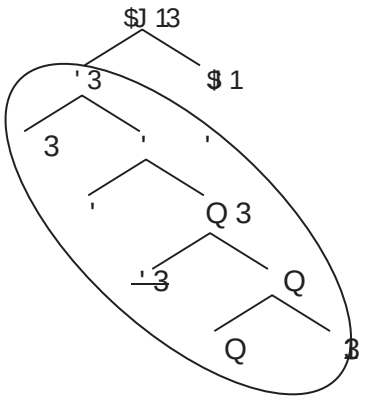


As is standard, we assume that D is a phase head, though we claim that in Hungarian it is a soft one, meaning that the nP is a soft domain, as indicated by the broken circle surrounding it in (23). As the possessor, in the specifier of nP, has no c-commanding competitor it is eligible for unmarked case. This assignment is not carried out at this point, the domain being soft. Instead, it is delayed until we reach the next hard domain. This happens when the AgN is merged with the DP. Being a hard phase head, its DP complement is a hard domain:



Lacking case at this point, the possessor is still visible and hence will be case assigned. As there is no c-commanding DP added to the structure, the possessor's situation remains unchanged and hence it is eligible for unmarked case assignment. The unmarked case it is assigned is the one defined for the domain that contains it, which is nominative. At this point, given that the nP and the DP have both been spelled out, the possessor is not available for movement and hence the nominative possessor always appears in the lower possessor position.

However, if the possessor had been moved to the specifier of DP before the merger of AgN, the following situation would pertain:



Now the possessor is in the hard DP domain, though it still has no c-commanding competitor and is still eligible for unmarked case, which can now be assigned. But this time the case involved is the one defined for complement of AgN, which as we have seen is dative. After case assignment, the possessor is invisible and therefore can play no further role in the case system. Any further extraction of the dative possessor will therefore not change its case.

11. Conclusion

We have argued that whether the case of a moved DP changes is dependent on both the case involved and the domain which contains it. Only if the case is unmarked and the domain soft can the case 'change'. However, the nature of case 'change' is, in reality, simply a hold on the assignment of unmarked case in a soft domain. Therefore, the relevant DPs remain visible, under the Invisibility Principle, and may be assigned a different case in a subsequent domain.

Having demonstrated that the claim in (21) accurately accounts for the interaction between movement and case in Hungarian possessor and dative subject extraction, we can highlight the conceptual advantages of the proposal. Note that (21) is a definition of a soft domain which is based on how case is assigned within it. The main property of a soft domain, that some of its content remains active after spell-out falls out directly from what we are taking to be a central part of case theory. The Invisibility Principle determines that a DP becomes inactive in the case system as soon as it is assigned case and hence delaying the assignment of unmarked case in some domains allows those DPs to remain active in subsequent domains, though not indefinitely so. Hard domains cause all contained DPs to become inactive because no DP can be left without case after these are spelled out. The overall picture is more consistent than just a simple statement that certain domains allow spelled out elements to remain active while others do not.

Furthermore, the fact that the distinction between soft and hard domains is defined specifically in terms of case theoretic considerations accounts for why the distinction is

relevant only for case phenomena and not movement. This better supports Baker's original proposal that the same domains are relevant for case and movement phenomena. The claim can be maintained in terms of the identity of the domains, though differences in how these domains limit phenomena follow from properties of the phenomena themselves.

Works Cited

An, Duk-Ho. 2014. "Genitive Case in Korean and Its Implications for Noun Phrase Structure."Journal of East Asian Linguistics í K W W S V G R L R U s10831-014-9123-9

Baker, Mark. 2015. Case: Its Principles and Parameters. Cambridge: Cambridge-University Press.

Chomsky, Noam. 2001. "Derivation by Phase." In Ken Hale: A Life in Language, edited by Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.

Dékány, Éva. 2018. "The Position of Case Markers Relative to Possessive Agreement." Natural Language & Linguistic Theory 36 (2): 365–400.

Den Dikken, Marcel. 1999. "On The Structural Representation of Possession and Agreement. The Case of (Anti-)Agreement in Hungarian Possessed Nominal Phrases." In Crossing Boundaries: Theoretical Advances in Central and Eastern European Languages, edited by István Kenesei, 137–178. Amsterdam: Benjamins.

Grimshaw, Jane. 2005. "Extended Projection." Words and Structure: 1–74. Stanford: CSLI.

Kishimoto, Hideki. 2013. "Case, Tense and Subject Raising in Japanese." Nanzan Linguistics í

É. Kiss, Katalin. 2014. "Ways of Licensing External Possessors in Hungarian." Acta Linguistica Hungarica í

Marantz, Alec. 1991. "Case and Licensing." ESCOL 91: Proceedings of The Eighth Eastern States Conference On Linguistics, edited by German Westphal, Benjamin Ao, D Q G + H H 5 D K N & K D H í & R O X P E X V 2 K L R 6 W D W

Nakamura, Hiroaki. 2002. "Double Subject, Double Nominative Object and Double Accusative Object Constructions in Japanese and Korean." Language, Information and Computation í

Newson, Mark, and Szécsényi, Krisztina. 2020. "Dative as the Unmarked Unmarked Case in Hungarian." In Linguistica Brunensia 68 (2), edited by Markéta Ziková D Q G 0 D U F L Q : J L H O í % U Q R 0 D V D U \ N 8 Q L Y H U V

Ntelitheos, Dimitris. 2002. "Possessor Extraction in Greek and the Left Periphery of DP." Ms. UCLA.

Szabolcsi, Anna. 1983. "The Possessor That Ran Away from Home." The Linguistic Review 3.1: 89–102.

Szabolcsi, Anna. 1994. "The Noun Phrase." In The Syntactic Structure of Hungarian: Syntax and Semantics Y R O H G L W H G E \ ) H U H Q F . L H I H U D C New York: Academic Press.

Tóth, Ildikó 2000. Inflected Infinitives in Hungarian. Phd Diss. Tilburg University.
Webelhuth, Gert. 1992. Principles and Parameters of Syntactic Saturation. New York: Oxford University Press.
Woolford, Ellen. 2006. “Lexical Case, Inherent Case, and Argument Structure.” Linguistic Inquiry í

Modal Existential Constructions in Hungarian

Zsolt Prohászka [a], Gábor Alberti [b], Anna Szeteli [c], and Judit Farkas [d]

University of Pécs, Department of Linguistics, Pécs, Hungary

[a]prohaszka.zsolt@pte.hu; [b]alberti.gabor@pte.hu; [c]anna.szeteli@gmail.com; [d]juttasusi@gmail.com

Abstract: The paper is devoted to a comprehensive description of the Hungarian modal existential wh-construction(s). These constructions consist of three main elements: an existential predicate, a pronoun, and a subordinate verb, which is either an infinitive or a subjunctive. Šimík (2011, 45) sorted languages into three groups based on the possible subordinate verb forms in modal existential wh-constructions. Hungarian belongs to the languages which can accept both infinitive and subjunctive verbs. This paper adds to this classification that the subjunctive type is significantly less acceptable in modal existential wh-constructions than the infinitival type but Transylvanian Hungarians significantly more readily use the subjunctive type than Hungarians within Hungary do. Then we turn to the theoretical investigation of the pronominal component of the construction, which is formally identical to interrogatives in Hungarian. We point out that it occupies a topic position within a [Topic(s)–Quantifer(s)–Focus] cycle-internal operator domain (É. Kiss 2002). In the final part of the paper, we have positioned the pronominal component of Hungarian modal existential wh-constructions, which is ultimately indefinite, in Haspelmath's (1997) semantic map of indefinites.

Keywords: modal existential wh-construction; infinitive; subjunctive; operators; indefinite pronouns

1. Introduction:

1.1 Modal Existential Wh-constructions in Hungarian

In this paper we examine Hungarian modal existential (MEC) and other constructions containing pronouns which are formally identical to interrogative wh-words in this language. Our main goals are to decide the information-structural status of such

pronouns—pronominal absolute stems—in MECs, and, in general, to present their high degree of multifunctionality in Hungarian, not discussed in Haspelmath's (1997, 291–292, Fig. 4.8.26) seminal typological description.

MECs consist of three main elements: an existential predicate, a pronoun, and a subordinate verb, which is either an infinitive or a subjunctive. Šimík (2011, 45) sorted languages into three groups based on the possible subordinate verb forms that can appear in MECs. The first group of languages only use the infinitive verb form. The second group consists of languages being able to use only the subjunctive in the given construction. Hungarian is in the third group: these languages can accept both infinitive and subjunctive verbs in MECs. Hungarian infinitives can also be inflected, therefore Hungarian MECs can appear in three forms regarding the subordinate verbs: MECs can host a bare/uninflected infinitive (1a); an inflected infinitive (1b), and a subjunctive subordinate verb (1c).

(1) (a) Van kit be-mutat-ni Ilinek.
is who.z|| preV-introduce- Ili.}z•
'There is someone to introduce to Ili.'

(b) Van kit be-mutat-n-om Ilinek.
is who. z|| preV-introduce- - Ili.}z•
'There is someone for me to introduce to Ili.'

(c) Van kit be-mutass-ak Ilinek.
is who. z|| pre-V-introduce- Œ€Ili.}z•
'There is someone for me to introduce to Ili.'

Constructions hosting an uninflected infinitive for the subordination are used in two ways: they can either stand together with an explicit, "named", subject (2a), or without it (2b). The (2a) type is understood the same way as the type with an inflected infinitive (1b), in that it is clear in both cases whom the statement expressed by the MEC regards. If the explicit subject is left out from the construction, as in (2b), MECs can be understood as universal statements, which are generally true for every possible or relevant person at the situation at hand.

(2) (a) Neked van mit olvasni. "named subject"
you.}z• is what.z|| read.,‡•
'There is something for you to read.'

(b) Van mit olvasni. “general subject”
is what. z | | read. , ‡ •
‘There is something (for everyone) to read.’

1.2 Other Forms of MECs

There are similar constructions in Hungarian which have been considered as MECs in previous studies (Šimík 2011, Lipták 2003).

Besides the existential van ‘be’, tud ‘can’ can also take the position of the main predicate (3a), presumably due to its inherently modal character. However, this type of MEC is restricted as it can only take the uninflected infinitive as subordinate verb, while other verb forms are unacceptable as the

(3) (a) Tudunk kit bemutatni Ilinek.
can. ‰ … who. z | | preV.introduce, ‡ • Ili. } z •
‘We can introduce someone to Ili.’

(b) (?)Van, akit bemutassunk Ilinek.
is a-who. z | | preV.introduceŒ Ž { ƒ ‰ }. z •
‘There is someone for us whom we can introduce to Ili.’

MECs can also be formed with relative pronouns instead of interrogatives (3b). Such MECs are also restricted as only the subjunctive verb form can occur in such constructions.

A common property of the two constructions presented in (3a-b) is that they cannot be used to express universal statements due to the obligatorily appearing verbal inflection (in different components of the MECs in question).

After this theoretical introduction to the rich world of modal existential constructions in Hungarian, in the following subsection we focus on the interchangeability of the competing subordinate verb forms in MECs, in order to get a more realistic picture on the factual distribution of the alternatives. Since the subordinate verbs cannot be altered in the constructions illustrated in this subsection, these have been left out from the investigation presented in 1.3 (but see Szabó and Prohászka 2021). We have restricted our attention to the construction types presented in 1.1, which are introduced by the existential predicate (some form of van ‘be’) and contain a pronoun that is formally identical to interrogative pronouns in Hungarian.

1.3 The Subjunctive as a Dispreferred Verb Form in MECs

To the best of our knowledge, previous literature – based on introspection – did not discuss any difference in the usage of the alternative subordinate verb forms (1a–c). This subsection is devoted to the verification of a hypothesis according to which the subjunctive version of MEC is dispreferred. We have conducted three experiments in which Google

forms were filled out by (non-linguist) native speakers of Hungarian and statistical anal yses were executed by SPSS. They had been asked for evaluating potential (written) sentences by means of the six-point scale, presented and defined in Table 1, which is to be regarded as the standard canonical scale of current generative linguistics, given that it has been applied in such seminal series as Comprehensive Grammar Resources for different languages (Broekhuis et al. 2012, Alberti and Laczkó 2018).

- Ž } € † ~ ‡ •	6 • z ‡ } z ‹ } , ‡ • ~ ‹ ‰ ‹ ~ • z • , ˆ ‡ ‘	& ˆ ‡ • ~ ‹ Œ , ˆ ‡ • ’ ‰ ~ Œ ‡ CGR				
		v1	v2	v3	v4	v5
*	unacceptable	–3	0	–5	–1	–1
*?	relatively acceptable compared to *	–2	1	–4		
??	intermediate or unclear status	0	2	0	0	
?	marked: not completely unacceptable, or disfavoured form	+1	3	+2		0
(?)	slightly marked, but probably acceptable	+2	4	+4	+1	
9	fully acceptable	+3	5	+5		+1

Table 1. The six-point scale of judgments proposed in CGR series and their conversion into different numerical scales

The distributions of grammaticality judgments concerning the three MEC minimal pairs provided in (4a–c) are presented in Figure 1. The pie charts clearly show that, relative to the practically fully acceptable infinitival versions, the subjunctive versions are deeply divisive for native speakers.

(4) (a) Van mit mondanod/mondj.
is what. z | | say. , ‡ • Œ €/say. Œ Ž { ƒ Œ €
‘There is something for you to say.’

(b) Van hol aludnod/aludj.
is where sleep. , ‡ • /sleep. Œ Ž { ƒ Œ €
‘There is a place for you to sleep.’

(c) Van kivel elmennem/elmenjek a boltba.
is who. z | | preV.go. , ‡ • /preV.go. Œ Ž { ƒ the shop. , … …
‘There is someone for me to go to the shop together with.’

Figure 1. Grammaticality judgments of native speakers on MECs with an infinitive / a subjunctive verb in the first experiment in 2018 (N=37; 8 males and 29 females, all native speakers of Hungarian within Hungary)

As shown in Figure 2 of corresponding MEC versions (where the corresponding confidence intervals do not overlap), the infinitival version is always significantly more acceptable than the subjunctive version. This holds true in spite of the fact that two of the three subjunctive versions are qualified as (not worse than) "slightly marked" on average (see Table 1 above), due to the high speaker-dependent standard variation, visualized by the pie charts in Figure 1.

1 As for the less acceptable, and most divisive, subjunctive version in (4a), its dispreferred status has presumably to do with the following specialty of Hungarian morphology. The subjunctive /imperative mood in Sg2 can be marked, in addition to the regular way with a Sg2 agreement V X v [mond-j-ál 'say- Œ Ž { ', without any explicit reference to any person and number (mond-j 'say- Œ Ž { *f* '), as in (4a)). Our conjecture is that there is some "division of labor" according to which a one-syllable variant is "too short" for a subordinate clause but it is preferably to use in the main clause of an imperative sentence.

Figure 2. Pairwise significant differences (95%) between members of the minimal pairs presented in Figure 1 according to the numerical scale numbered as v1 in Table 1

Figure 3 below is devoted to illustrating a major methodological question. How can we avoid the loss of information coming from using statistically the system of the six grammatical judgments given in Table 1 merely as ordinal variable, instead of interpreting it as an interval variable (Field 2013, 8–9), with appropriately chosen distances between the six values? On the basis of our linguistic experiences, we propose that the conversion given in Table 1 labelled as v1 from the ordered six values into numbers. This conversion follows two basic rules: i. the extreme points of the scale of judgments are associated with a positive number and its negative counterpart (namely, +3 and –3), ii. "intermediate or unclear status" is mapped onto zero, in harmony with this definition. At first glance, it seems to be a disadvantage of conversion v1 that no judgment is mapped onto –1; but this can also be regarded as the true reflection of the linguistic practice according to which we have had more grades for "almost good" than for "almost bad".

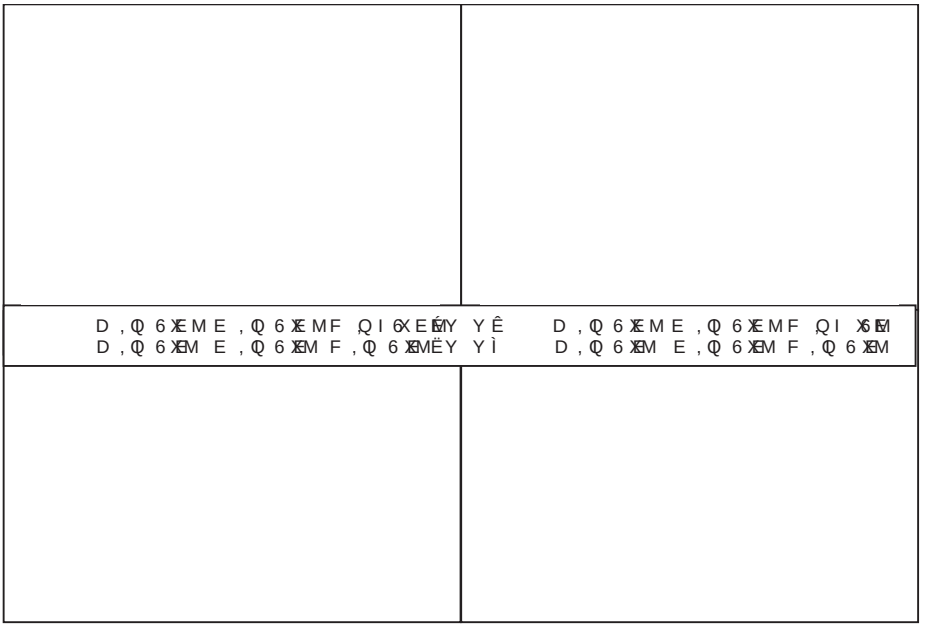


Figure 3. Pairwise significant differences (95%) between members of the minimal pairs presented in Figure 1 according to the numerical scales numbered as v2, v3 (first row), v4 and v5 (second row) in Table 1

By exhibiting the four sets of error bar charts (Field 2013, 106) in Figure 3, calculated based on four alternative ways of "intervalization" of the originally only ordered judgments, we intend to convince the reader that the choice of conversion, as far as it is reasonable from a linguistic point of view, does not radically influence the statistical output. As shown by the pairwise non-overlapping confidence intervals in all four parts of Figure 3, infinitive variants still prove to be significantly more acceptable than the corresponding subjunctive counterparts; even if certain judgments are not distinguished any more during the conversion process, as in the case of v4 and v5.

If the MEC-internal subject is named, as in (2a) and (5), due to ALSO-quantification, for instance, which blocks the otherwise almost obligatory pro, infinitival versions are still significantly more acceptable than their subjunctive counterparts in the minimal pairs, see Figure 4.

(5) (a) Nekem is van hova mennem/menjek.
For I. } z• also is where go. , ‡ • /g Œ € Œ € Ž { ƒ Œ €
'There is some place for me, too, to go to.'

(a') Én is van hova menjek.
I also is where go. Œ Ž { ƒ Œ €
'There is some place for me, too, to go to.'

(a") Mi is van hova menjünk.
we also is where go. Œ Ž { ƒ Œ €
'There is some place for us, too, to go to.'

(b) Neked is van kivel beszélned/beszélj.
For you. } z • also is who. , ‡ Œ speak. , ‡ • Œ €/speak. Œ Ž { ƒ Œ €
'There is someone for you, too, to speak to.'

(c) Neki is van P L W O tartania/tartson.
For he. } z • also is what. z { …fear. , ‡ • Œ €/fear. Œ Ž { ƒ Œ €
'There is something for him, too, to be afraid of.'

(5a)/Inf_{Agr} (5a)/Subj (5b)/Inf_{Agr} (5b)/Subj (5c)/Inf_{Agr} (5c)/Subj

Figure 4. Pairwise significant differences (95%) between members of the minimal pairs presented in (5a, b, c) according to the numerical scale numbered as v1 in Table 1. The experiment was carried out in 2020 (N=37; 4 males and 33 females, all native speakers of Hungarian within Hungary)

Since it is not obvious whether the named subject in subjunctive MEC variants is dative (due to the copular component *van* 'is', cf. the possessive construction *nekem van* 'I have (got)…',) or nominative case marked (as in basic subjunctive clauses in Hungarian), a third experiment addressed this question, see Figure 5. In this experiment our attention has also been extended to a Transylvanian dialect of Hungarian, which is claimed to use the subjunctive mood more frequently than the dialects of Hungarian spoken within the territory of Hungary (É. Kiss 2009, 214).

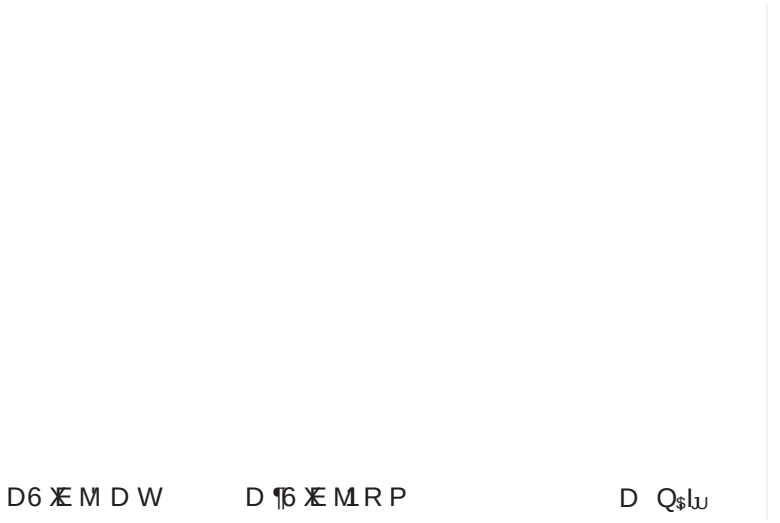


Figure 5. For speakers within Hungary, as shown by the bars on the left-hand side in the attached pairs, there is no significant difference between the subjunctive MEC variants given in (5a–a') while the infinitival variant in (5a) is significantly more acceptable than both subjunctive variants. The experiment was carried out in 2021 (N=46; 21 native speakers of Hungarian within Hungary, 25 native speakers of a Transylvanian dialect of Hungarian).

The two subjunctive variants, with different case marked named subjects, do not show significant difference either within Hungary or in Transylvania. The infinitival counterpart, however, can be claimed to be significantly more acceptable only within Hungary, as shown by the corresponding overlapping confidence intervals for the Transylvanian dialect in Figure 5.

Figure 6 below presents further evidence for the more preferred status of the subjunctive version of MEC in the Transylvanian dialect of Hungarian, relative to the standard version of Hungarian. There are, however, analogous examples in the case of which the difference between the Transylvanian data and the data from Hungary only "almost" reach the level of 95% significance.

Figure 6. Sentence (5a"), representative of the subjunctive version of MEC with a native case marked named subject, is significantly more acceptable in Transylvania (see the bar to the right) than in Hungary, according to the third, 2021 experiment

To summarize, the infinitival version of MEC in Hungarian tends to be significantly more acceptable than the subjunctive version, which we consider to serve as a contribution to the relevant literature. We have also pointed out that there are dialectal variations in this respect: Transylvanians' bias to subjunctive also manifests itself in the area of modal existential constructions.

2. The Operator Position of the Wh-pronoun in the Hungarian [Topic* Quant* Focus] Operator Sequence (É. Kiss 2002) in MEC

2.1 Focus?

The syntactic position of the pronominal component in MECs is debated. The standard generative literature automatically considers the pronominal component in MECs to function as a (narrow) focus, see example (11a-b) in É. Kiss (2002, 202–203). Lipták (2006, 6) also claims that this wh-item occupies a (narrow) focus position, despite the preverb-verb order, which can be seen in (1a–c) in 1.1.

The series of examples in (7) provides the background. First, it is a characteristic feature of Hungarian narrow focus that it triggers the verb-preverb order in finite constructions (7a). Second, in wh-questions, the wh-word occupies the focus position; in these cases, the preverb-verb order is unacceptable either, as illustrated in (7b). In the case of infinitival constructions containing a narrow focus, however, the verb-preverb order is optional (7c); which makes it impossible to immediately refuse Lipták's hypothesis.

(7) (a) Csak Petit hívtam be / *be-hívtam a megbeszélésre.
only Peti.z|| called.Œ€ preV/preV-called.Œ€ the meeting.ŒŽ{
'It is only Peti whom I called into the meeting.'

(b) Kit hívtál be / *be-hívtál a megbeszélésre?
who.z|| called.Œ€ preV/ preV-called.Œ€ the meeting.ŒŽ{
'Who did you call into the meeting?'

(c) Szeretném Csak Petit hív-ni be / be-hívni az értekezletre.
like.|^‡}.1Œ€ Only P.z|| call-,‡• preV preV-call-,‡• the meeting.ŒŽ{
'I would like to invite only Peti to the meeting.'

In an earlier paper (Prohászka et al. 2020), based on statistically analyzed tests we presented that the verb–preverb order immediately following the wh-pronoun of MECs is not optional at all but highly refused by native speakers ('*?', see Table 1 in 1.3), as shown in Figure 7 below. This is a strong argument for refusing the hypothesis that the pronominal component of MECs occupies a (narrow) focus position.

(8a)/Inf (8b)/Inf (8a)/Inf_{Agr} (8b)/Inf_{Agr} (8c)/Inf (8d)/Inf (8c)/Inf_{Agr} (8d)/Inf_{Agr}

Figure 7. Significant difference between fully acceptable MECs with the verb-verb order and unacceptable MEC variants with a verb-preverb order (Prohászka et al. 2020, 66)

(8) (a) Ilivel van mit megbeszélni megbeszélnem.
Ili. , ‡ Œ is what. z | | preV.discuss. , ‡ • preV.discuss. , ‡ • Œ €
‘There is something (for me) to discuss with Ili.’

(b) Van mit elmesélni / elmesélnem Marinak.
is what. z | | preV.tell. , ‡ • preV.tell., ‡ • Œ € Mari. } z •
‘There is something (for me) to tell Mari.’

(c) *? Ilivel van mit beszélni / beszélnem meg.
Ili. , ‡ Œ is what. z | | discuss. , ‡ • discuss. , ‡ • preV
Intended meaning: ‘There is something (for me) to discuss with Ili.’

(d) *? Van mit mesélni / mesélnem el Marinak.
is what. z | | tell. , ‡ • tell. , ‡ • Œ preV Mari. } z •
Intended meaning: ‘There is something (for me) to tell Mari.’

Another argument against the focus-view is that more than one wh-element can be present in Hungarian MECs, as in (9); and the preverb is emphasized, which does not hold, for instance, for the preverb in (7c) in a real narrow-focus construction.

(9) Vankit kinek bemutatnom.
is who. z | | who. } z • preV.introduce, ‡ • Œ €
‘There is at least one person whom I can introduce to at least one person.’

2.2 Quantifier or Topic?

,Q D SRWHQWLDO e .LVV VW\OH û ù ù û >7RSLF 4XH
domain, as discussed above, the wh-pronoun cannot serve as (even a) focus, since it cannot trigger inversion; whereas a real identificational focus does trigger inversion even within a MEC, at least as an option, see Prohászka et al. (2020, 67). The interpretation of the wh-pronoun in MECs does not support the hypothesis according to which it is an ALSO-quantifier or an EACH-quantifier, either. The meaning of this element is closer to valaki ‘someone’. Such pronouns are typically claimed to serve as topics, see É. Kiss (2002, 106, ex. (7a)), for instance (10).

(10) [$_{TopP}$ Valaki szerencsére [$_{AspP}$ meghívta Jánost]]
Somebody luckily preV.invited. Œ € János. z | |
‘Somebody luckily invited John.’

We propose thus that the pronoun of MECs is similar to valaki-type pronouns in that they both occupy a topic position in the operator sequence. The operator domain of MECs presented in our study is as follows:

(11) (a) Van kit$_{Topic}$ kinek$_{Topic}$ bemutatnom /?bemutassak.
is ki. z | | ki. } z• preV.introduce. , ‡ • Œ preV.introduce. Œ Ž { ƒ Œ €
'There is at least one person whom I can introduce to at least one person.'

(b) Van mit$_{Topic}$ csak veled$_{Focus}$ (?)megbeszélnem ?[beszélnem meg].
is mi. z | | only you. , ‡ Œ preV.talk. , ‡ • Œ talk. , ‡ • Œ € preV
'There is at least one topic which I can discuss only with you.'

The topic-interpretation of the existentially bound pronoun of MECs thus explains possible word-order variants. However, as the MEC-pronoun is always non-specific, it violates the criterion of topichood proposed by É. Kiss (2002, 11):

> The formal features of topic:
> A topic constituent must be [+referential] and [+specific].

This contradiction can be resolved in two ways. The first solution is to consider the topic of MECs as exceptional: the pronouns in MECs are exceptional topics in that they do not need to meet the referentiality criterion, which applies for topics in general.

We propose a more elegant solution. Our claim is that the function of a topic expression consists of pure anchoring to referents in the scope in which the given expression can be found, while additional conditions depend on the character of the scope in question. In this theory, the "scope" of topics that belong to finite verbs is the discourse itself, in which the given sentence is to be interpreted; hence, these topics should be anchored to (salient) discourse referents, as in (10). This anchoring provides a [+referential] and [+specific] interpretation to the given topic.

In MECs, however, the scope in question is an existential stratum. In this environment, the topic is interpreted as an existentially bound expression (x.P(x)). Therefore, the topic of MECs is not "exceptional", but is determined by the accommodating scope typical of MECs.

2.3 Commenting on Surányi's (2005) Proposals concerning the Interpretations of Hungarian MEC-Pronouns

Surányi (2005), based on (allegedly ambiguous) sentences like the one presented in (12), claims that the interpretation of a MEC-pronoun can be both existential and universal. This subsection is devoted to arguing against the latter option.

(12) (Most aztán) (Jánosnak) van kinek mit adnia.
now then János.}z• is ki.}z• what. z | | give. , ‡ • Œ €
(a) ‘Now John has something to give to everyone.’
y (PERSON(y) o x (THING(x) & GIVE(j,y,x)))
(b) ‘John has things to give to people.’
x (THING(x) & y (PERSON(y) & GIVE(j,y,x))

On the basis of the assumed alternative readings, his conclusion is that the MEC-pronoun of a MEC is quantified either by sitting in DistP (12a), universally, or by being bound by the existential quantifier of the copular MEC-verb (12b). We agree with the latter proposal: we consider this element to be the straightforward MEC-specific existential topic proposed in 2.2.

As for reading (12a), we claim that the construction in (12) is not ambiguous with alternative meanings (12a) and (12b). Instead, we consider this to be vague, with (12a) being a very special case within the general meaning given in (12b): ‘at least one pair’ might mean ‘several pairs’. This means that the semantic content given in (12b) is basically the intersection of that presented in (12a) and our world knowledge (‘in certain situations one should give some present to each person in a relevant set’), whereas in the case of (11a) in 2.2, for instance, world knowledge does not produce a (12a)-type reading: ’for each relevant person there is someone whom I can introduce him or her[2].

3. The Place of MEC-pronouns in the System of Hungarian Indefinites [3]

The fact that MEC-pronouns formally coincide with bare-interrogative pronouns is of great importance in the light of Haspelmath’s (1997, 2003) extensive typological research into systems of indefinites. According to Haspelmath (1997, 27), bare-interrogative pronouns play a crucial role in the system of indefinites, whose universal systematization, in the form of semantic maps, belongs to Haspelmath’s distinguished aims in his seminal typological book on indefinites across languages of the world:

---

2 Thus we argue against the stance that a MEC-pronoun can be associated with the type of universal interpretation which, for instance, the accusative case marked wh-word in (17c) in 3.3 can be associated with (in a sequence of wh-words). Surányi’s proposal obviously comes from the observation that the sentence Kinek mit adott János? ‘who.}z• what. z | | gave János?’ is, indeed, (and not an interrogative pronoun): The hearer is asked for revealing for each relevant person what János gave to him/her

3 Special thanks are due to Hans-Martin Gärtner for raising our attention to Haspelmath’s method of constructing semantic maps for pronoun types.

"[I]ndefinites are a marked category relative to interrogatives … (Moravcsik 1969,77) … there is a universal asymmetric markedness relation such that indefinite pronouns are usually more marked than and derived from interrogative pronouns." "Indefinite pronouns are … derived by conversion from interrogative pronouns."[4]

He proposes a universal semantic map based on data from 40+100 languages (in which linked pragmasemantic functions are expressed by the same phonetic forms in certain languages) with 9 universal functions, see Figure 8. The crucial typological generalization can be formulated as follows: if two functions are associated with the same pronominal forms in a language, the functions between these two along the lines can also be associated with the given pronominal form in the given language.

Figure 8. An implicational map for functions of indefiniteness pronoun series as proposed by Haspelmath (1997, 64)

In Hungarian (Haspelmath 1997, 291–292), four series of indefinite pronouns are claimed to cover the universal functional space with the nine functions, as presented in Figure 9. A few examples are valaki 'somebody', bárki/akárki 'anybody', and senki 'nobody'.[5]

4 It is also emphasized at the end of the book what central research question bare-interrogative-indefinite pronominal series raise: "Some puzzles that bare-interrogative-indefinite pronominal series present are that they are rather heterogeneous in their morphological structure, that they are rarely fully productive, that they are rarely fully grammaticalized, and that they rarely have a single function" (Haspelmath 1997, 291).

5 The partial né-series, e.g. néhány 'a few', is mentioned in Haspelmath's book (1997, 291–292) but is not shown on the map. As for the representatives of the four complete sets of indefinite pronominal series with the [+WH] bare-interrogative form, which can be replaced with other interrogative forms. Valaki and akár are quite transparently derived from van 'be' and akar 'want', respectively. Senki is ne(m) is+nem 'also+not' (Szabolcsi 2018, 240) where the also-like component has an emphasizing contribution. These three etymological sources for indefinite pronominal series, a system-like alternative to akár, coincides with the connective bár '(al)though'.

Figure 9. Haspelmath's semantic map for Hungarian indefinites

Due to a potential indefinite interpretation of the German jeder 'each (person)', Haspelmath (1997, 155) expands the general semantic map to (non-specific) universals. We claim that in Hungarian we can differentiate altogether 10 different pragmasemantic functions associated with the bare-interrogative indefinite form in the further-expanded version of this map, as presented in Figure 10. The bare-interrogative form thus can be characterized by an outstandingly high level of hidden multifunctionality; and it is also surprising how diversely the functions are associated with different operator characters in the É. Kiss-style (2002) Topic / Quant / Focus trichotomy, discussed in 2.2.

Figure 10. The semantic map we propose for Hungarian indifinites; the ten (!) functions of bare-interrogative pronouns are linked by thick lines

In the final part of the paper, we present the different (Haspelmathian) functions which can be associated with bare-interrogative pronominal forms (such as *kicsoda* 'who', *mi* 'what', *hol* 'where', and their suffixed variants).

3.1 Three Related Functions of MEC-pronouns in Haspelmath's System: Specific Known / Specific Unknown / Irrealis Nonspecific

In the case of specific known, the speaker knows what (in this case: which place) the indefinite pronoun refers to, as illustrated by the potential continuation.

(13) Van hol aludnod; …
is • ˆ … sleep. , ‡ • Œ €
‘There is a place for you to sleep; (you will sleep at Donners’.)’

The situation of specific unknown is somewhat different from the previous one, as although someone knows what (or here: which place) the indefinite pronoun refers to, it is not necessarily the speaker.

(14) Van hol aludnod; …
is • ˆ … sleep. , ‡ • Œ €
‘There is a place for you to sleep; (Mari has mentioned the family’s name, but I have forgotten).’

In the type of irrealis nonspecific, it is not known what the indefinite pronoun refers to.

(15) Van hol aludnod; …
is • ˆ … sleep. , ‡ • Œ €
‘There is a place for you to sleep; (there are six of you and the local families have offered places for eight).’

To sum up, the constructions discussed in subsection 3.1, three (related) functions of MEC-pronouns have been differentiated from the Haspelmathian perspective, expressed “otherwise” in Hungarian by the vala-series of indefinite pronouns (e.g. valaki ‘someone’, valami ‘something’, valahol ‘somewhere’, valahogy ‘somehow’). As vala- is related to van ‘be’, it is a straightforward explanation for the fact that MEC-pronouns “remain” bare interrogatives that in MECs they are typically immediately preceded by the verb van itself. The modal existential construction can be regarded as the place of the Hungarian grammar where the source of the vala-series of indefinites has preserved. In respect of operator character (É. Kiss 2002), the pronoun types illustrated in (13–15) all function as topics, in the sense elaborated in 2.2 (according to which a topic is not necessarily specific).

3.2 Interrogative

The series of examples in (16) is devoted to the presentation of what is generally held to be the basic use of the absolute pronominal stem, that is, the interrogative use. Haspelmath (1997, 2003) does not classify interrogatives as indefinite pronouns; nevertheless, it is worth considering that in Hungarian an interrogative object ab ovo triggers the indefinite conjugation, as shown in (16a), but interrogatives can also be used as definite expressions, witnessed by the conjugation presented in (16b). This may legitimize an approach according to which certain interrogatives belong to indefinites while members of the complementary subtype are definites.

(16) (a) Kit engedtek / *engedték be? …
KI. z | | let. %$_{IndefObj}$ let. %$_{DefObj}$ preV
‘Who was allowed to go in?’

(b) Kidet engedték / *engedtek be?
KI. Œ €$_{Poss}$ z | | let. %$_{DefObj}$ let. %$_{IndefObj}$ preV
‘Who, of your family members or acquaintances, was allowed to go in? Your sister or your mother?’

(c) Ki mindenkit engedtetek /*engedtétek be?
KI everyone. z | | let. %$_{IndefObj}$ let. %$_{DefObj}$ preV
‘Who (presumably many people) did you allow to come in?’

To make the picture complete, we illustrate in (16c) a complex interrogative construction, obligatorily triggering the indefinite conjugation, in which an interrogative pronoun is combined with a universal pronoun (Bartos 2020). In respect of operator character (É. Kiss 2002), the pronoun type illustrated in (16) functions as a narrow focus, witnessed by the “inverse” order between preverbs and verb stems.

3.3 Universal

7 Z R I D F W V D U H L O O X V W U D W H G L Q) L U V W L Q
V S H F L ¿ F D D V Z H O O D V G H ¿ Q L W H V S H F L ¿ F E X
Second, in a list of pronominal absolute stems, only the last member is interpreted as an interrogative pronoun with the preceding ones to be interpreted as universal
T X D Q W L ¿ H U V W e ± V H H W K H W U D Q V O D W L R Q L Q
to use a regular universal pronoun is forbidden in the construction in question. It is
D O V R L O O X V W U D W H G L Q F W K D W W K L V kal L Q G R I
ovo W U L J J H U V W K H L Q G H ¿ Q L W H W F \ S H O H M S X U R Q R L Q D Z K L Y O H
permitted in the construction). In respect of operator character (É. Kiss 2002), thus, the
S U R Q R X Q W \ S H L O O X V W U D W H G L Q F I X Q F W L R Q V D

(17) (a) Mindenkit beengedtek /*beengedték.
every.KI. z | | preV.let. %$_{IndefObj}$ preV.let. %$_{DefObj}$
‘Who (presumably many people) did you allow to come in?’

(b) Mindegyik lányt beengedték /*beengedtek.
all girl. z | | preV.let. %$_{DefObj}$ preV.let. %$_{IndefObj}$
‘Both girls were allowed to go in.’

(c) Kit hova engedtek /*engedték be?
KI. z | | hova let. ‰$_{IndefObj}$ let. ‰$_{DefObj}$ preV
'Who were allowed to go in where?' [The addressee is expected to give information on each relevant person in the given respect (where were they allowed to go in).]

(c') *Mindenkit hova engedtek /engedték be?
every.KI. z | | hova let. ‰$_{IndefObj}$ let. ‰$_{DefObj}$ preV
Intended meaning: (16c)

3.4 Different Ways of Multiplying the Reference of Bare Interrogatives

The first way of multiplying the reference of bare interrogatives yields multiple partitive constructions, which can be observed in Hungarian with the bare-interrogative form; (18) illustrates two different realizations of the type.

(18) (a) Ki pénzre vár, ki egy levélre,
KI money. Œ Ž { wait. Œ € KI a letter. Œ Ž { …
ki újságokra.
KI newspaper. ‰ … Œ Ž { …
'Some (people) are waiting for money, some for a letter, some for newspapers.'

(b) Elszaladtak, ki merre látott.
prev.ran. ‰ ..KI MERRE saw. Œ €
'They dispersed, some in some direction, some in other directions (i.e. different people in different directions).'

The bare-interrogative form can also take part in reduplication

(19) Menjen táncba ki-ki köztünk D] jegyesével!
go. Œ Ž { ƒ dance. , ..KI-KI between. ‰ the his/her fiancé.‰ ˆ Œ Œ Œ € , ‡ Œ
'Everyone among us should go to dance with their own fiancé.'

In respect of operator character (É. Kiss 2002), the pronoun types illustrated in (17–19) function as specific topics.

3.5 Relative Pronoun

The bare-interrogative pronominal form in the place of the standard relative pronoun (e.g. aki 'that who') provides archaic flavor (20). As for its potential place in the Hungarian system of indefinites, the relevant piece of information is presented in (20a): the pronoun (also in its archaic bare-interrogative disguise) in the subordinate clause triggers

the indefinite conjugation while the corresponding pronoun in the main clause the definite conjugation. In respect of operator character, relative pronouns function as another type of operator than Topic, Quantifier or Focus; they function as a subordinating operator (É. Kiss 2002, 243).

(20) archaic Kit beengedtek, azt mi is beengedtük volna.
KI. z | | preV.let. %$_{\text{IndefObj}}$ that. z | | we also preV.let. %$_{\text{DefObj}}$ be. | ˆ ‡ }
'Those who were allowed to go in would have been allowed to come in by us too.'

3.6 Indirect and Direct Negation

We can meet in Hungarian the type of indirect negation in rhetorical questions. The bare-interrogative can practically be interpreted as 'nobody' in the given context; in respect of operator character (É. Kiss 2002), however, it functions as a (narrow) focus.

(21) Ki engedne be ilyen gyanús alakot?!
KI let. | ˆ ‡ } Œ € preV such suspicious figure. z | |
'Who would allow to go in someone as suspicious as this person?'

It is also archaic, but quite acceptable in many constructions, to substitute the non-human bare-interrogative pronominal form for the standard negative universal pronoun *semmi* 'nothing', which can be regarded as the direct type of negation. The pronoun type illustrated in (22) functions as a universal quantifier.

(22) archaic Mit sem ér.
MI. z | | neither be_worth. Œ €
'It is worth nothing.'

4. Conclusion

The paper has been devoted to a comprehensive description of the Hungarian modal existential construction(s).

Section 1 presents the results of three experiments. These have been carefully analyzed by statistical methods after converting the canonical ordinal system of grammaticality judgments applied in modern generative linguistics into an interval variable which is ideal input to pointing out significant differences in different areas. We have verified that the subjunctive type of MEC is significantly less acceptable than the infinitival type but Transylvanian Hungarians use the subjunctive type significantly more readily than Hungarians within Hungary do.

Section 2 is devoted to theoretical argumentation. We claim that the pronoun of modal existential wh-constructions, which is formally identical to interrogatives, occupies a topic position, and nothing else (see subsection 2.3). The given topic position

is considered within the [Topic(s)–Quantifer(s)–Focus] MEC-internal operator domain, proposed on the basis of the standard hypotheses concerning the left periphery of the Hungarian sentence structure (É. Kiss 2002). One might say that this is an exceptional topic, considering its non-referential character (cf. É. Kiss 2002, 11). Instead of this view, we have argued that the common characteristic of topics lies in their anchoring feature, and the superordinate scope will specify the cycle-specific nature of different subtypes of topic.

Section 3 discusses MEC-pronouns, which formally coincide with bare interrogatives, from a Haspelmathian (1997, 2003) typological perspective. We have pointed out that this common pronominal form shows extraordinary multifunctionality in Haspelmath's (1997) universal semantic map of indefinite pronouns with its 8 present-day (+2 archaic) functions. The MEC-pronoun, this hidden indefinite, occupies a much more expanded domain in the Hungarian map (Haspelmath 1997, 291–292) than the well known *vala/né, akár/bár* and *se(n)*series (corresponding to the *some*-, *any*- and *no*-series in English, respectively). Its extraordinary multifunctionality is also seen in the fact that the ten functions are associated with Topic, Quantifier, Focus as well as Subordinating character in the É. Kiss-style (2002) operator system.

Acknowledgement

The paper is supported by The Syntax of Hybrid Categories project (NKFI FK-128518).

Works Cited

Alberti, Gábor, and Tibor Laczkó, eds. 2018. Syntax of Hungarian – Nouns and Noun phrases. Amsterdam: Amsterdam University Press.

Broekhuis, Hans, Evelien Keizer, and Marcel Den Dikken. 2012. Syntax of Dutch – Nouns and Noun Phrases. Amsterdam: Amsterdam University Press.

%DUWRV +XED ³0LPLQGHQW WXGXQN D PLPLQ
we know about the 'what everything' type questions?]" In Általános nyelvészeti tanulmányok XXXII.: Újabb eredmények a grammatikaelmélet, nyelvtörténet és XUDOLV]WLND N|UpEO [New results in grammar theory, history of languag tics], edited by Éva Dékány, Tamás Halm, and Balázs Surányi, 249–264. Budapest: Akadémiai Kiadó.

e .LVV .DWDOLQ ³$]HJ\V]HU& PRQGDW V]HUNH
clause]" In Strukturális magyar nyelvtan [A Structural Grammar of Hungarian], edited by Ferenc Kiefer, 79–177. Budapest: Akadémiai Kiadó.

É. Kiss, Katalin. 2002. The Syntax of Hungarian. Cambridge: Cambridge University Press.

Field, Andy. 2013. Introduction to Statistics using IBM SPSS. London: Sage Publ.

Haspelmath, Martin. 1997. Indefinite Pronouns. Oxford: Oxford University Press.

Haspelmath, Martin. 2003. “The Geometry of Grammatical Meaning: Semantic Maps and Cross-linguistic Comparison.” In The New Psychology of Language, Vol. 2., edited by Michael Tomasello, 211–242. Mahwah, NJ: Lawrence Erlbaum Associates.

Lipták, Anikó. 2003. Hungarian Modal Existential Wh-constructions. Manuscript. Leiden: University of Leiden.

Moravcsik, Edith A. 1969. “Determination.” Working Papers in Language Universals 1, 64–98. Stanford: Stanford University.

Prohászka, Zsolt, Judit Farkas, Gábor Alberti, and Anna Szeteli. 2020. “The Position of Preverbs in Hungarian as an Indicator of Verbalness.” In Exploring Variation in Linguistic Patterns, edited by Karolina Drabikowska and Anna Prazmovska, 55–77. Lublin: Wydawnictwo KUL.

Šimík, Radek. 2011. Modal Existential Wh-constructions. Utrecht: LOT.

Szabó, Veronika and Zsolt Prohászka. 2021. “Hungarian Modal Existential Wh-constructions in Written Language Use – Diachronic Perspectives.” Paper presented at International Conference on the Structure of Hungarian 15, Pécs, Hungary, 25–26 August 2021. https://icsh15.netlify.app/icsh15_abstracts/ICSH15_paper_Szabo_Prohaszka.pdf.

Szabolcsi, Anna. 2018. “Strict and Non-Strict Negative Concord in Hungarian: A Unified Analysis.” In Boundaries Crossed, at the Interfaces of Morphosytax, Phonology, Pragmatics and Semantics, edited by Huba Bartos, Marcel den Dikken, Zoltán Bánréti, and Tamás Váradi. 227–242. Cham: Springer.

Surányi, Balázs. 2005. “Indefinite Wh-pronouns and the Modal Existential Wh-construction in Hungarian.” Paper presented at International Conference on the Structure of Hungarian 7, Veszprém, 29–31 May 2005. Handout: http://www.nytud.hu/icsh7/absz/Indefinite%20wh-pronouns.pdf.

VSO-VOS Alternations in Kaqchikel

Seid Tvica

Georg August University, Göttingen, Germany

seid.tvica@uni-goettingen.de

Abstract: In this paper I argue against the uniform head movement analysis for Mayan languages, as proposed by Clemens and Coon (2018). According to their analysis VOS orders with definite objects are predicted not to exist. However, exactly such orders have been attested in at least one variety of Kaqchikel (cf. Broadwell 2000). Furthermore, Kaqchikel exhibits free VSO-VOS alternations with both arguments being indefinite. To account for the VSO-VOS alternations, I argue that the two orders are derived in syntax. I consider two different analyses. In one I employ uniform leftward movement with exclusively specifier-initial orders, while in the other I consider both leftward and rightward movement as well as specifier-initial and specifier-final orders. The second analysis appears to be superior as it adequately accounts for the absence of adverbial intervention as well as for the ungrammaticality of orders in which definite arguments precede the indefinite.

Keywords: syntax; (anti)symmetry; head movement; rightward movement

1. Introduction

In this paper I argue against the uniform head movement analysis for Mayan languages, as proposed by Clemens and Coon (2018) (henceforth, C&C) in which both VSO and VOS orders are derived via verb movement to the clause-initial position. Their account that the surface VOS orders follow from, (i) postsyntactic reordering of arguments, (ii) right-side subject topicalization and (iii) heavy NP shift do not coherently account for discourse neutral VSO-VOS alternations in Kaqchikel, a Mayan language spoken in Guatemala. Crucially, Kaqchikel exhibits VOS orders with definite objects, which

under C&C's analysis are predicted not to exist, as none of the three conditions are met. To account for the VSO-VOS alternations in Kaqchikel, I argue that the two orders are derived in syntax. I present two different analyses. In one analysis, I employ an anti-symmetric syntax (involving uniform leftward specifiers and leftward movement) in which either the subject DP or the VP moves to Spec,TP position, yielding the respective orders. In the other analysis, I employ a symmetric syntax (involving both leftward and rightward movement) in which arguments surface as rightward multiple vP specifiers, which can be freely ordered at PF, yielding VOS or VSO. Either analysis accounts for the basic word order facts in Kaqchikel, as documented in Broadwell (2000), although they make different predictions. In particular, the symmetric (rightward specifier) analysis appears to be superior for Kaqchikel as it accounts for the following two additional facts in the language that cannot be captured with the antisymmetric (uniform leftward movement) analysis.

First, the antisymmetric analysis requires stipulation that definite arguments cannot precede the indefinite (*<DEF,INDEF>), suggesting that in an analysis where all syntactic objects move to the left, the definite argument cannot be higher than the indefinite (*DEF>INDEF, where '>' stands for 'higher than'). This is unexpected and runs counter to the observation that DEF>INDEF is attested many languages. In contrast, under the symmetric analysis *<DEF,INDEF> naturally follows as DEF moves over INDEF to a higher position on the right side, and thus, DEF>INDEF is the expected hierarchical order.

Second, the antisymmetric analysis inevitably leads to the generation of structures in which the verb surfaces outside of vP. This turns out to be problematic as adverbials that typically adjoin to vP cannot intervene between the verb and its arguments in either VOS or VSO orders. This also runs counter to cross-linguistic tendency of a number of different adverbials attaching low in the structure, which are typically assumed to be vP-adjoined. In contrast, under the symmetric analysis, both arguments remain vP-internal and therefore ipso facto vP-adjoining adverbs are possible, and yet (correctly) cannot intervene.

This paper is structured as follows. In section 2, I discuss C&C's analysis on VSO-VOS alternations in Mayan languages and show that the facts from Kaqchikel challenge the uniform head movement analysis along with the C&C's proposals that seemingly violating superficial orders are post-syntactically derived. In section 3, I develop an analysis involving antisymmetric syntax which exclusively involves leftward movement, and show that while such an analysis can account for the VSO-VOS alternations, it fails to adequately account for the aforementioned *DEF<INDEF restriction as well as the absence of adverbial intervention. To account for these facts, in section 4, I argue that syntax is underlyingly symmetric (as traditionally assumed) allowing rightward specifiers within vP and (short distance) rightward movement. This adequately captures the Kaqchikel facts. Section 5 concludes.

2. VOS in Mayan

2.1 Clemens and Coon (2018)

In many Mayan languages, VOS orders readily occur when objects are bare NPs, but are unavailable with DP objects, as illustrated in (1) and (2).

(1) (a) Tyi y-il-ä x'ixik wiñik
PFV a3-see- woman man
'The man saw the woman.' (Vázquez Álvarez 2011, 21)

(b) Tyi i-kuch-u si' aj-Maria
PFV a3-carry- wood CLF-Maria
'Maria carried this wood.(Coon 2010, 355)

(2) (a) *Tyi y-il-ä jiñi x'ixik wiñik
PFV a3-see- DET woman man
intended: 'The man saw the woman.'

(b) *Tyi i-kuch-u ili si' aj-Maria
PFV a3-carry- DEM wood CLF-Maria
intended: 'Maria carried this wood.' (Coon 2010, 355)

Determiners, demonstratives, and proper names cannot appear as objects in VOS orders. C&C observe that the presence of D layer material on objects correlates with the word order as only objects that lack D material on top of the nominal can appear in VOS orders (Clemens and Coon 2018, 247). if the object is a DP, the resulting order must be VSO:

(3) Tyi i-kuch-u aj-Maria ili si'
PFV a3-carry- CLF-Maria DEM wood
intended: 'Maria carried this wood.'

To account for the VOS-VSO alternations in Mayan, C&C consider a number of potential analyses for this. Right-side specifiers (a la Aissen 1992) are rejected as this presents complications for rigid VSO languages like Qanjob'al and Mam. Although C&C do not elaborate on why this is the case, presumably, this may be due to the fact that subjects appear to the left of the object and therefore the assumption that the subject is in a spec position is unlikely. Similarly, vP-fronting analysis by Coon (2010) is rejected because such analysis is difficult to extend to languages with rigid VSO orders.

In order to have a uniform account of the verb-initial orders, C&C argue that VOS is derived from VSO in three different ways (i) postsyntactic restructuring of bare NP objects, (ii) 2 heavy-NP shift (iii) right-side topics. Following verb (head) movement to the clause-initial position, objects that lack D^0 layer material postsyntactically move to yield VOS orders in order to maintain prosodic constituency. Evidence for this analysis comes from Cho'l, in which, according to C&C, the prosodic boundaries in VOS and VSO orders differ in that there is no prosodic boundary between the verb and the object in VOS orders. In contrast, in VSO orders, the prosodic boundaries are H Y L G H Q W E H W Z H H Q D O O H O H P H Q W V D V L O O X V W U D boundaries.

(4) Prosodic phrasing of VSO and VOS clauses in Ch'ol

D 9 ¥ 6 ¥ 2 ¥

E 9 2 ¥ 6 ¥

(Clemens and Coon 2018, 252)

Based on the prosodic boundaries and the presence of the D^0 layer material, C&C contend that a nominal with a D^0 layer is a phase, whereas the object that lacks the D^0 layer (i.e. an NP object) is not a phase. As such, NP objects undergo leftward movement and surface adjacent to the verb with which they form a prosodic constituent. This analysis predicts that "naturally occurring examples of VOS in cases where the object is a DP and the subject is neither a topic nor a heavy NP would constitute counterevidence to [C&C's] claims" (Clemens and Coon 2018, 274).

Although in their analysis C&C primarily focus on the facts in Ch'ol, they contend that their analysis holds for all Mayan languages. However, we show below that Kaqchikel has an additional set of data that suggests that the VOS-VSO alternations are already available in syntax.[1] In addition to Kaqchikel, Tz'utujil, Poqomam, and Poqomchii' may also exhibit orders that challenge the notion that the variation arises postsyntactically.[2] In the remainder of this section, I focus on Kaqchikel data that challenge C&C's post-syntactic account. While I draw data primarily from Broadwell (2000) in my analysis, the reader should be made aware that there are distinct varieties of Kaqchikel for which a different set of facts (and therefore the analysis) may hold.

2.2 Kaqchikel

According to Broadwell (2000), transitive clauses in Kaqchikel exhibit VOS, VSO, and SVO orders, as illustrated in (6a), (6b), and (6c) respectively:

1 See Douglas, Ranero, and Sheehan (2017) who also argue for a syntactic analysis for word order alternations in Mayan.

2 Thanks to Michelle Sheehan (p.c.) for pointing out Poqomam and Poqomochii'.

(6) (a) x-u-b'a ri tz'i' ri me's
COMPL-3.SG.ERG-bite the dog the cat
'The dog bit the cat.'

(b) x-u-b'a ri me's ri tz'i'
COMPL-3.SG.ERG-bite the cat the dog
'The dog bit the cat.'

(c) ri tz'i' x-u-b'a ri me's
the dog COMPL-3.SG.ERG-bite the cat
'The dog bit the cat.' (Broadwell 2000)

Verb-initial orders show ambiguity if the subject and the object have "equal degrees of definiteness" (cf. Broadwell 2000). That is, either of the argument DPs following the verb can either be the subject or the object of the verb if both arguments are definite or if both arguments are indefinite (cf. 7):

(7) (a) x-r-oqotaj ri tz'i' ri me's
COMPL-3.SG.ERG-chase the dog the cat
'The dog chased the cat.'
'The cat chased the dog.'

(b) x-r-oqotaj jun tz'i' jun me's
COMPL-3.SG.ERG-chase a dog a cat
'A dog chased a cat.'
'A cat chased a dog.'

The ambiguous data in (7) run counter the predictions made by C&C's analysis. Firstly, these cases are ambiguous in their interpretation (p.c. George Aaron Broadwell), suggesting that there cannot be any prosodic differences between the two readings. Secondly, the definite objects do appear in VOS readings. And thirdly, both arguments are light NPs, which per definition means that that these cases cannot involve any instances of heavy NP shift.

In addition to Kaqchikel, these facts may also hold in other Mayan languages. For example, for Tz'utujil, C&C adopt their data from Dayley (1985), who documents that VSO orders are absent. However, Duncan (2003) explicitly points out that, according to his consultants, the VSO orders are available patterning with Kaqchikel. Compare (7) and (8).

(8) x-uu-tz’et jun tz’i’ jun miix
INC-3.SG.ERG-see a dog a cat
‘A cat saw a dog.’
‘A dog saw a cat.’ (Duncan 2003, 169)

Furthermore, heavy objects in Tz’utujil can appear in either VSO or VOS, suggesting that the obligatory heavy-NP shift is not at play:

(9) (a) x-uu-tz’ub’-aj derja Aa Xwaan ru-chi Ta Mari’y
COMPL-3.ERG-kiss- the Mr Juan 3.ERG-cheek Miss Maria
‘Juan kissed Maria on the cheek.’

(b) x-uu-tz’ub’-aj der ru-chi Ta Mari’y ja Aa Xwaan
COMPL-3.ERG-kiss- 3.ERG-cheek Miss Maria the Mr Juan
‘Juan kissed Maria on the cheek.’

Regarding the contexts in Kaqchikel with different degrees of definiteness in subject and objects, if one of the argument DPs is definite and the other one indefinite, then the definite argument must be the subject and it must follow the object yielding VOS orders. In the following example the definite DP must follow the indefinite DP, and it must be the subject of the verb:

(10) (a) x-r-oqotaj jun me’s ri tz’i’
COMPL-3.SG.ERG-chase a cat the dog
‘A dog chased a cat.’
‘*A cat chased a dog.’

(b) *x-r-oqotaj ri tz’i’ jun me’s
COMPL-3.SG.ERG-chase the dog a cat

If the subject is indefinite and object definite then verb-initial orders are ungrammatical and the surface order must be SVO. In this case, actor focus (AF) morphology appears as a suffix on the verb (cf. 11a), and cannot be left out (cf. 11b).

(11) (a) Jun tz’i’ x-b’a’-o ri a Juan
A dog COMPL-bite-AF the CL Juan
‘A dog bit Juan.’

(b) ?*Jun tz’i’ x-u-b’a’ ri a Juan
A dog COMPL-3.SG.ERG-bite the CL Juan

The summary of possible word orders depending on the definiteness of subject and object is given in the following table:

(12)

Subject-object definiteness and word orc		
S_{indef}	O_{indef}	VOS, VSO, SVO
S_{def}	O_{def}	VOS, VSO, SVO
S_{indef}	O_{def}	SVO
S_{def}	O_{indef}	VOS

Given (12) we can make a relevant generalization that the analysis of the clause structure should capture. Namely, the definite arguments in Kaqchikel cannot precede the indefinite. I refer to this as the *<DEF,INDEF> restriction.

Regarding the placement of adverbials, verb-initial and the SVO orders differ in that in verb-initial adverbials cannot intervene between the verb and its arguments (cf. 13). However, the intervention effects are attested in SVO orders (cf. 14).

(13) (a) iwir x-r-oqotaj ri tz'i' ri me's AdvV SO
yesterday COMPL-3.ERG-chase the dog the cat
(b) x-r-oqotaj iwir ri tz'i' rime's *VAdvSO
(c) x-r-oqotaj ri tz'i' iwir ri me's *VSAdvO
(d) x-r-oqotaj ri tz'i' ri me's iwir ?VSOAdv

(14) (a) iwir ri tz'i' x-r-oqotaj ri me's AdvSVO
yesterday the dog COMPL-3.ERG-chase the cat
(b) ri tz'i' iwir x-r-oqotaj ri me's SAdvVO
(c) ri tz'i' x-r-oqotaj iwir ri me's SVAdvO
(d) ri tz'i' x-r-oqotaj ri me's iwir ?SVOAdv

Given (13) and (14) there are (at least) three possible ways of deriving the structures: (i) SVO orders are derived from verb-initial orders, (ii) verb-initial orders are derived from SVO orders, (iii) verb-initial orders and SVO orders are derived independently. Since adverbs are much more flexible in SVO, it stands to reason that SVO orders are more flexible and involve additional movement operations. It is for this reason that in my analysis below, I adopt the option (i) and leave aside the option (ii). The option (iii) is a priori undesirable, and I leave it aside (though see Broadwell 2000 who treats them as independent derivations).

In the next section I discuss an analysis that involves antisymmetric syntax and illustrate that such an analysis comes with (ad hoc) stipulations. Subsequently, in section 4, I demonstrate that these stipulations are not required under a symmetric analysis.

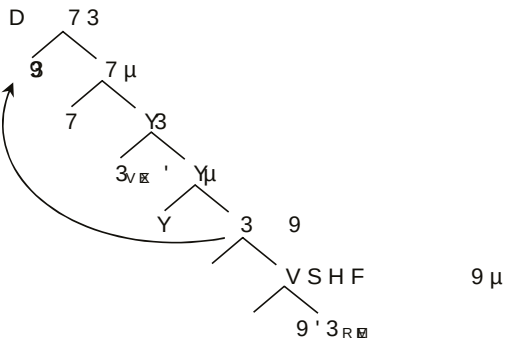
3. Leftward Movement Analysis

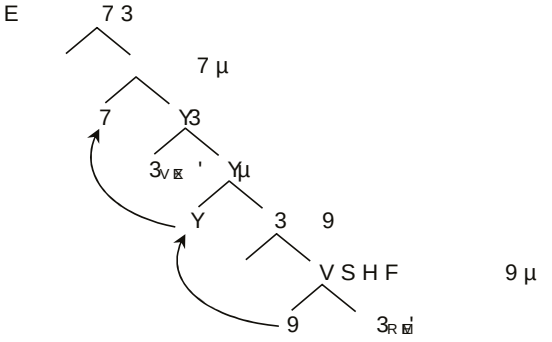
Under the assumption that verb-initial orders are default in Kaqchikel, we can surmise that since the verb and the object are not adjacent in VSO orders it follows that under the standard assumptions the displacements must be taking place. It is well understood that verbs combine with objects first before they combine with subjects, as a generalization which Baker (2010) dubs as Verb-Object Constraint (VOC). Since the verb and object are not adjacent in VSO, it follows that either the verb or the XP containing the verb moves over the subject to the left. Alternatively, the object may move to the right, over the subject in the rightward specifier position. However, since antisymmetric syntax requires uniform leftward movement as well as specifiers on the left side, this option cannot be considered here, but I return to it in section 4.

Given the prerequisites of antisymmetric syntax, both VSO and VOS orders must be derived. Two potential approaches exist in the literature that can derive the orders, namely, head movement analysis and VP- (and VP-remnant) movement analysis. Under head movement analysis the verb moves over the subject deriving VSO from SVO. However, to derive VOS orders, we must stipulate object shift crossing the subject. This would require a postulation of a strong feature that optionally triggers object shift (recall that verb-initial orders with arguments of equal degree of definiteness can optionally read as VSO or VOS). This rules out a uniform head movement analysis.

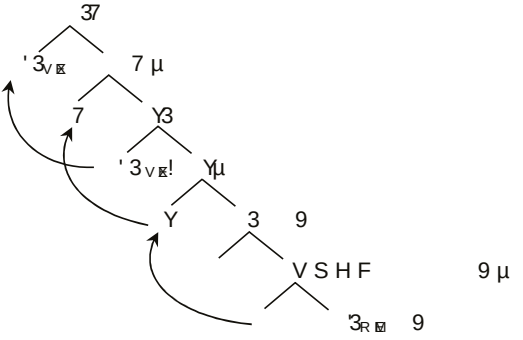
Under a VP-movement analysis, the VP containing the verb and its object move to a position in front of the subject, which derives VOS orders. However, in order to derive VSO orders, we run into the same problem (as with the head movement analysis) of having to postulate a strong feature that optionally triggers object shift out of VP, before the VP fronts. Consequently, we are left with the option of having both VP-movement and verb movement. That is, either VP moves to the front along with the direct object yielding VOS or verb alone moves to the front yielding VSO, as illustrated in (15a) and (15b), respectively.

[illegible] are ambiguous (Aaron Broadwell p.c.). If correct then the option of deriving either order likely [illegible] triggers either v^0 or VP movement should then be syntactic rather than a discourse feature, as the reviewer proposes.



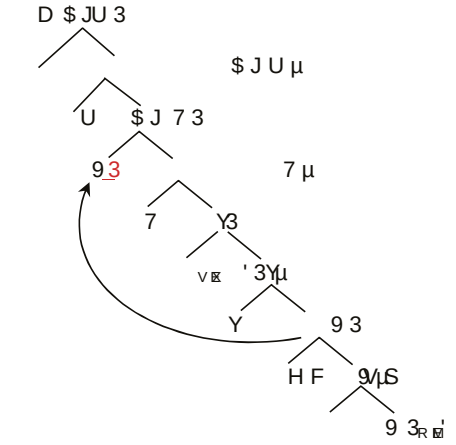


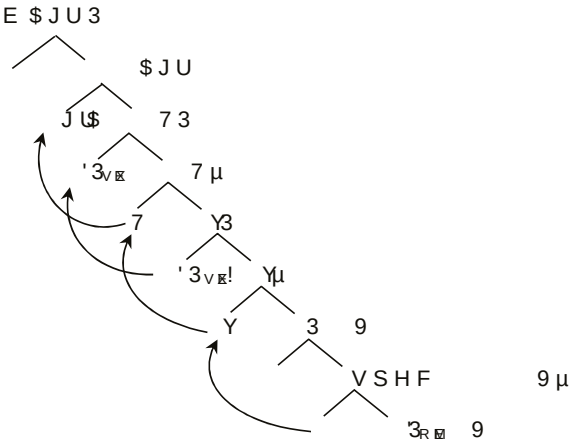
In (15), I assume that the projection to which either VP or the verb moves is TP. Furthermore, I assume that the TP projection bears an EPP feature that must be checked either by VP, or in the case of verb movement by the subject which raises to spec,TP in (15b) yielding the structure in (16).



Subsequently, the verb raises to Agr (cf. 17) deriving the correct word order. Since Kaqchikel is morphologically rich involving aspectual morphology as well as agreement

morphology that appear on the verbal stem, verb movement to the higher position can be triggered by any of these higher inflectional morphemes. For example, there is a long standing view that head movement is often triggered by affixal requirements of higher functional heads, such as agreement morphology (cf. Bobaljik and Thráinsson 1998; Rohrbacher 1994, 1999; Koeneman 2000; Koeneman and Zeijlstra 2014; Tvica 2017). Thus, (15a-b) are derived to yield the following two structures, respectively:





In (17a) where the VP containing the verb raises to spec,TP, the verb and the affix in the Agr^0 position appear adjacent at PF which allows the affix to appear on the verb, whereas in (17b) the verb raises to Agr^0 where the affix attaches to the verb. Importantly, when arguments are of equal degree of definiteness the syntax generates either (17a) or (17b), allowing for ambiguous interpretation of verb-initial orders, i.e. either VOS or VSO, respectively.

However, when subjects are definite and object indefinite, only (17a) is possible. In contrast, when subjects are indefinite and objects definite, the order must be SVO which requires subject movement to clause initial position (spec,AgrP or higher). Importantly, the subject movement to spec,AgrP must be available in both (17a) and (17b), as SVO orders allow adverbial intervention in all positions. Concretely, if the subject movement to spec,AgrP is available only in (17a) then we cannot derive SVAdvO orders, which are attested. If, however, subject to spec,AgrP is available only in (17b) then we cannot derive SAdvVO orders.

Nevertheless, even with both structures employed to generate all possible adverb placements in SVO orders, problems still arise. The generation of SAdvVO orders with the subject movement to spec,AgrP, requires that adverbs must be adjoined to TP. If this is correct, then it is not clear why adverbs cannot intervene between the verb and the subject in VSO orders, as generated with the structures in (17b). More generally, there is an additional problem for (17), namely, why can't adverbs adjoin to VP in Kaqchikel? This certainly comes as a surprise as low adjunction of certain types of adverbs is common in many other languages.

Although the analysis correctly accounts for the VOS-VSO alternations, (17) must involve the *DEF<INDEF restriction, suggesting that in the antisymmetric syntax the indefinite arguments must be higher than the definite. This is a stipulatory stipulation, which is a priori undesirable. More importantly, it runs counter to crosslinguistic findings, according to which it is typically the case that the definite arguments raise higher a position, as in many Germanic languages.

In the next section, I develop a symmetric analysis that accounts for both the lack of adverbial intervention and for the surprising restriction on the linear order of arguments with unequal degree of definiteness (i.e. *<DEF,INDEF>).

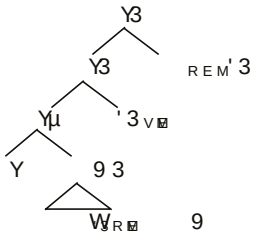
4. Rightward Movement Analysis

4.1 Spec-Final Orders

In order to mitigate the ordering constraint *DEF<INDEF let us suppose that all the specifiers are adjoined to the right. In Kaqchikel, there is some independent evidence for this assumption from possessive phrases. Standardly the possessor and the possessee, e.g.in English, are assumed to occupy specifier and complement positions, respectively. This corresponds to the left-right <possessor, possessee> linear order. In Kaqchikel, however, the linear order is opposite (cf. 18), already superficially suggesting the rightwardness of specifiers.

(18) n-u-kanoj r-ixjayil a Manuel rija'
CON-3.SG.ERG look:for 3.SG.ERG wife Manuel s/he
'Manuel's wife is looking for him.'

Furthermore, to allow the optionality of the postverbal argument orders in (7), let us assume that subjects and objects of equal degrees of definiteness surface as specifiers of a single head. There is substantial literature on multiple specifiers in which it has often been observed that in certain cases the order of specifiers can alternate F I & K R P V N \ 5 L F K D U G V % R ã N R Y L ü - L P p Q H) H U Q i Q G H] - L P p Q H]) H U Q i Q G H] D Q G , V V H Y H U in the free order of wh-words in multiple-wh-fronting languages, such as Serbo-Croatian D Q G 5 X V V L D Q Z K H U H W K H \ F D Q D S S H D U D V P X O W L S O This idea can be extended to vP, which would project multiple specifiers that can host both arguments of a transitive verb in Kaqchikel. With the assumption that multiple specifiers are on the right (as has been proposed by Aissen 1992), we can postulate that the Kaqchikel vP projects the surface structure in (19).



Here, the object DP appears at the outer specifier, while the subject at the inner specifier of vP. Crucially, given the VOC, the object DP must be base-generated as the sister of V, suggesting that it must undergo movement to the outer specifier of vP.

4.2 Rightward Movement

The structure in (19) involves a critical deviation from what has standardly been observed, namely that movement is predominantly to the left. Contrary to this, the object DP in (19) moves to the right. Here I take a position that movement to the right is possible, and although it is rare, I assume that there are no syntax-internal principles that forbid it (contra Kayne 1994). Rather, the apparent scarcity of rightward movement may follow from grammar-external factors such as parsing, as argued for by Ackema and Neeleman (2002) and Abels and Neeleman (2012), and even modality, as rightward movement has been observed in sign language (cf. Cecchetto et al. 2009).

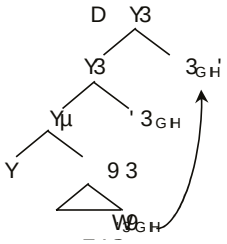
4.3 Basic Clause Structure in Kaqchikel

With the assumption that syntax is symmetric, involving instances of rightward movement, as given in (19), there are (at least) two hypotheses with respect to the order of postverbal arguments (of equal degrees of definiteness) that yield either VOS or VSO orders in Kaqchikel.

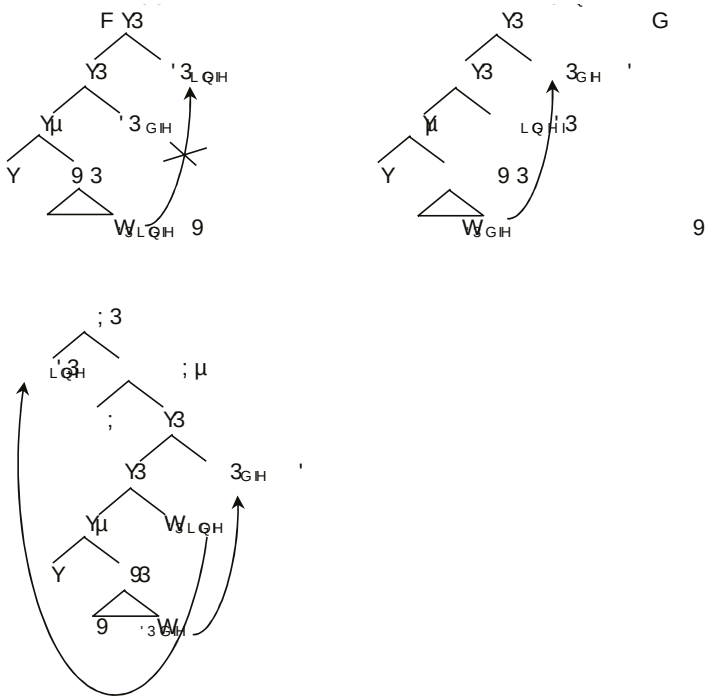
A. Objects must raise to the outer spec,vP, which allows free linearization of subjects and objects at spell-out.
B. Objects optionally raise to the outer spec,vP, which yields either order derivationally.[4]

Let us consider both hypotheses. Under hypothesis A, object shift is uniform in all derivations. (20a) derives both VOS_{def} and $VS_{def}O_{def}$ orders, while (20b) derives the indefinite counterparts. Structurally in both cases the object is syntactically higher, however, since both arguments are multiple specifiers of the same head (the v0) they can be linearized in either order. This patterns with the free ordering of wh-words in multiple-wh-fronting languages, as well as in multiple-topic-fronting languages, where WKH PXOWLSOH VSHFLILHUV RI & FDQ DSSHDU LQ HL ZLWK PXOWLSOH VSHFLILHUV RI 73 DV SURSRVHG E\

Regarding the clauses with arguments of different degrees of definiteness, (20c) derives $VO_{indef}S_{def}$. Here, the indefinite object surfaces in situ and cannot move higher over the definite subject, in accordance with the *<DEF,INDEF> restriction. Lastly, the $VS_{indef}O_{def}$ orders can in principle be derived as (20d) illustrates. However, this configuration is ungrammatical, as only SVO orders are available when subjects are indefinite and objects definite; In such configurations, the subject must escape to the clause-initial position (cf. 21).



4 A reviewer suggests that there should be some mechanisms that allow for rightward PRYHPHQWV 8QGHU WKLV DFFRXQW WKH OLQHDUL]DWLRQ must follow heads in Kaqchikel vPs, as opposed to for example the LCA, which cannot linearize VSHFL¿HUV WR WKH ULJKW



Under hypothesis B, object optionally raises to the outer spec of vP in both (20a) and (20b), yielding both VOS and VSO, where both arguments are definite in (20a) and both indefinite in (20b). In (20c) object cannot raise to the outer spec,vP as this violates the *<DEF,INDEF> restriction. (20d) can in principle be derived, however configurations with indefinite subjects and definite objects must yield SVO orders, so the subject must escape to the clause initial position.

For both hypotheses A and B scenarios, it seems to be the case that multiple specifiers of v^0 cannot be occupied by arguments of different degrees of definiteness, as object shift cannot take place in (20c). However, in (20d) it might take place, but we cannot detect it as, in this context, the indefinite subject must appear before the verb.

The crucial difference between the A and B scenarios is that under A, object always moves to the outer spec,vP, whereas under B, object movement is largely optional and it seems to have the same problem that we have shown for the antisymmetric analysis in section 3, where we would have to stipulate an optional strong feature that triggers object shift. In addition, the B scenario employs optional movement of definite objects, which raises the question of why (20d) is not possible? For A, which we adopt and contend to be superior, it can be argued that it is precisely because the object obligatorily raises over the subject to the outer spec,vP that the indefinite subject must escape to the clause-initial position in (21). In addition, the idea that definite objects optionally raise is challenged

by some empirical evidence pertaining to the Principle C effects, which I discuss next, suggesting that definite objects must move to a position from which they c-command the subject, such as the outer spec,vP.

4.4 Principle C

In Kaqchikel verb-initial orders, a pronoun in the object position can bind an R-expression inside the subject. According to Broadwell (2000) when the possessor of the subject is antecedent to a following pronoun as in (22a), SVO order is obligatory (cf. 22b).

(22) (a) n-u-kanoj r-ixjayil a Manuel rija'
CON-3.SG.ERG look:for 3.SG.ERG wife CL Manuel s/he
*'Manuel's$_i$ wife is looking for him$_i$.'
'Manuel's$_i$ wife is looking for him$_j$.'

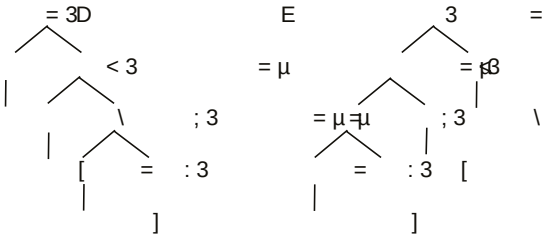
(b) r-ixjayil a Manuel | … Manuel n-u-kanoj rija'
3.SG.ERG wife CON-3.SG.ERG look:for s/he
'Manuel's$_i$ wife is looking for him$_j$.'

In addition to SVO orders in which subjects are indefinite and objects definite (cf. 12), in (22b) both subject and object are definite. If definite objects optionally raise over the subject, as is the case under scenario B, then the object could remain in a position from which it cannot c-command the subject at spec,vP. In other words, a VOS order should be available, but this according to Broadwell (2000) is not possible.

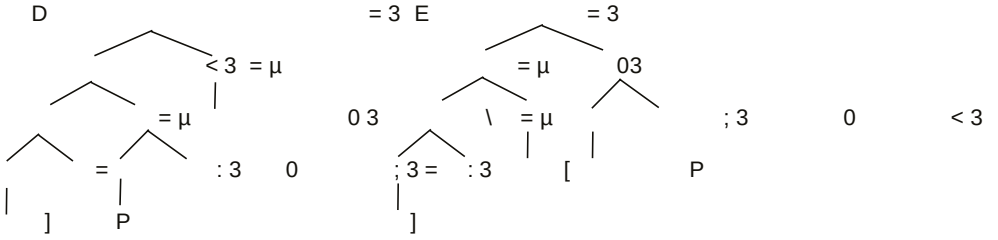
Since both arguments wind up in multiple specifiers the question arises how the c-command relations are established between the two specifiers of the same head. Note that Abels and Neeleman (2012) abstractly show that asymmetric c-command can hold between two specifiers of the same phrase, which suggests that if both arguments are specifiers of the same head, then there can be a c-command relation between the two, which could lead to binding effects (see also Guimarães 2008). This suggests that under the analysis in (20) the object in the outer specifier of vP must c-command the subject in the inner specifier of vP. In terms of geometric relations this indeed is this case. However, under Kayne's (1994) notion of c-command as given in (31) the two specifiers symmetrically c-command each other.

(23) X c-commands Y iff X and Y are categories and X excludes Y and every category that dominates X dominates Y. (Kayne 1994, 16)
X excludes Y if no segment of X dominates Y. (Chomsky 1986, 9)

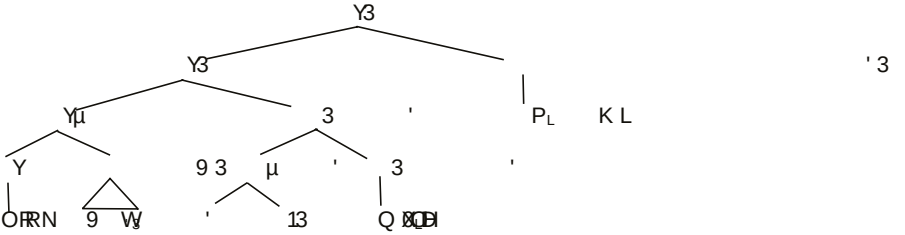
Abstractly, (23) stipulates that in the following multiple-specifier structures, neither YP nor XP asymmetrically c-command each other:



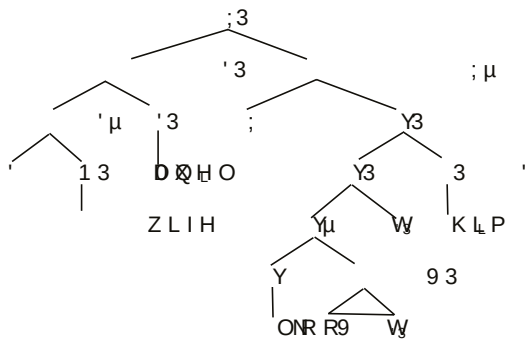
But, in (25a) YP asymmetrically c-commands XP, and in (25b) XP asymmetrically c-commands YP.



(25a) is exactly the structure of (22a), which is represented in (26), where the object pronoun (DP1) in the clause-final position asymmetrically c-commands the subject (DP2):



Since (22a) is ungrammatical under a bound reading, the apparent principle C violation in (26) is mitigated by the subject movement to the clause-initial position as represented in (27). Here, the subject moves to a projection XP, the nature of which is not relevant for the purposes here. However, as Kaqchikel productively uses actor focus morphology in SVO orders. The AF morpheme, typically appearing as a suffix on the verb, could be hosted by X, X^0 triggering verb movement.



DP movement to the left (rather than to the right) can potentially be attributed to processing limitations (cf. Ackema and Neeleman 2002). The movement of the subject in (27) is likely a long-distance movement, due to potentially other projections between vP and XP (such as TP and AspP). Hence, it is not surprising that it is to the left as cross-linguistically there is practically no evidence of long-distance rightward movement. Ackema and Neeleman (2002) suggest that the absence of long-distance rightward movement is due to processing limitations.

Since binding relations hold between arguments occupying multiple specifiers, it is reasonable to expect that this is something that we find in multiple specifiers elsewhere. Indeed, this is precisely what has been observed in the order of the multiply-fronted topics in Romance, but then from the higher position on the left side (cf. López 2009), a mirror image of what we see in Kaqchikel. A pronoun in the leftmost fronted topic in Spanish binds into the nominal within the subsequent fronted topic, violating Principle C. This is demonstrated in (28c) which is ungrammatical under the relevant reading.

(28) (a) Cada$_i$ niño puso su$_i$ chaqueta en el armario
Each child put-3SG.SELF his/her coat in the closet
'Each child put his/her coat in the closet.'

(b) Cada$_i$niño su$_i$chaqueta la puso en el armario

(c) */??Su$_i$chaqueta cada$_i$niño la puso en el armario

Under this analysis the attested placement of adverbs is straightforwardly accounted for. If both arguments remain vP-internal with multiple specifiers being available in vP, then it is expected that adverbs (standardly assumed to be vP-adjoined) cannot intervene between the verb and its arguments in VOS and VSO orders. The relevant assumption here is that the multiple specifiers of vP are freely ordered at PF (under hypothesis A), allowing both VOS and VSO. The crucial part of the assumption is that the free ordering

holds only between the elements in the specifiers contained within vP, without the inclusion of adjuncts, or else the system would incorrectly generate VSAdVO and VOAdvO orders.

The added benefit of the analysis is that the much greater freedom of adverbs in SVO orders follows straightforwardly. As discussed above, SVO orders are assumed to be derived from verb-initial orders and involve subsequent subject movement to the clause initial position as well as verb movement to the head of the projection that hosts the AF morphology, which appears as a suffix on the verb. The extraction of the subject and the verb to vP-external position yields adverbial intervention effects that are not attested in verb-initial orders.

5. Concluding Remarks

In this paper I have argued that a uniform head movement analysis as proposed by Clemens and Coon (2018) does not straightforwardly account for verb-initial orders in Kaqchikel, as the VSO-VOS alternation appears to be syntactic, rather than post-syntactic. I have therefore discussed two potential ways of deriving the VSO-VOS alternation, involving antisymmetric (with uniform leftward movement) and symmetric analyses. There appear to be two empirical hurdles for the antisymmetric analysis: (i) it is unclear why adverbial intervention is blocked in verb-initial orders, and (ii) the ad hoc stipulation that definite arguments cannot precede the indefinite runs counter to crosslinguistic findings, as definite-before-indefinite orders are widely attested. Neither of these issues arise under the symmetric analysis, suggesting that traditional syntactic theories that allow for mirror-image structures might be on the right track.

Funding Acknowledgement

This research has been supported by the German Research Foundation (Deutsche Forschungsgemeinschaft (DFG), grant #STE 958/10-1 / ZE 1040/6-1.

Works Cited

Abels, Klaus, and Ad Neeleman. 2012. “Linear Asymmetries and the LCA.” Syntax 15 (1): 25–74.

Ackema, Peter, and Ad Neeleman. 2002. “Effects of Short-Term Storage in Processing Rightward Movement.” In Storage and Computation in the Language Faculty, edited by Sieb Nooteboom, Fred Weerman and Frank Wijnen, 219–256. Springer, Dordrecht.

Aissen, J. 1992. “Topic and Focus in Mayan.” Language 68: 43–80.

Baker, Mark C. 2010. “Formal Generative Typology.” In The Oxford Handbook of Linguistic Analysis, 2nd ed., edited by Bernd Heine and Heiko Narrog, 285–312. Oxford: Oxford University Press.

Bobaljik, Jonathan David, and Höskuldur Thráinsson. 1998. “Two Heads Aren’t Always Better than One.” Syntax 1 (1): 37–71.

% R ã N R Y L ü ä H O M N R ³ 2 Q Linguistic Inquiry 33 (3): 351–380. W L Q J
Broadwell, George Aaron. 2000. “Word Order and Markedness in Kaqchikel.” Proceed
ings of the LFG ’00 Conference, edited by Miriam Butt and Tracy Holloway King,
44–62. Stanford University: CSLI Publications.
Cecchetto, Carlo, Carlo Geraci and Sandro Zucchi. 2009. “Another Way to Mark
Syntactic Dependencies: The Case for Right-Peripheral Specifiers in Sign
Languages.” Language 85 (2): 278–320.
Chomsky, Noam. 1986. Barriers. Vol. 13. Cambridge, MA, MIT Press.
Chomsky, Noam. 1995. The Minimalist Program. Cambridge, MA: MIT Press.
Clemens, Lauren, and Jessica Coon. 2018. “Deriving Verb-Initial Word Order in Mayan.”
Language 94 (2): 237–280.
Coon, Jessica. 2010. “VOS as Predicate Fronting in Chol.” Lingua 120 (2): 345–378.
Dayley, Jon Philip. 1985. Tz’utujil Grammar. Vol. 107. Berkeley: University of-Cali
fornia Press.
Douglas, Jamie, Rodrigo Ranero, and Michelle Sheehan. 2017. “Two Kinds of Syntactic
Ergativity in Mayan.” Proceedings of GLOW in Asia XI, edited by Michael
Yoshitaka Erlewine, 41–56. Cambridge, MA: MIT Press.
Duncan, Lachlan. 2003. “The Syntactic Structure of Tz’utujil Maya.” In Proceedings
of the LFG ’03 Conference, edited by Miriam Butt and Tracy Holloway King,
164–183. Stanford University: CSLI Publications.
Guimarães, Maximiliano. 2008. “A Note on the Strong Generative Capacity of Standard
Antisymmetry Theory.” Snippets 18: 5–7.
- L P pQ H]) H U Q iQ G H] È Q J H O / ³ 2 Q W K H 2 U G H U R I
Feature Inheritance.” ' L O E L O L P $ U D ú W (Journal of Linguistic
Research) 2011 (1): 5–32.
- L P pQ H] $ D Q G 6 ø ú V H Y H U ³ 0 X O W L S O H 7 R S L F
Feature Inheritance.” Paper presented at the Mediterranean Syntax Meeting 3,
8–10 October 2010, University of Athens.
- L P pQ H]) H U Q iQ G H] È Q J H O / D Q G 6 H O oX N ø ú V H Y H U
Tronting: Intervention of Focus and Binding”. In Current Issues in Generative
Linguistics: Syntax, Semantics and Phonology, edited by - R D Q Q D % á D V] F] D N
8–25. Wroclaw: Center for General and Comparative Linguistics (CGCL).
Kayne, Richard S. 1994. The Antisymmetry of Syntax. Cambridge, MA: MIT Press.
Koeneman, Olaf. 2000. “The Flexible Nature of Verb Movement.” PhD diss., Utrecht
University.
Koeneman, Olaf, and Hedde Zeijlstra. 2014. “The Rich Agreement Hypothesis Reha
bilitated.” Linguistic Inquiry 45 (4): 571–615.
López, Luis. 2009. “Ranking the Linear Correspondence Axiom.” Linguistic Inquiry
40 (2): 239–276.

Neeleman, Ad, and Amanda Payne. 2020. “PP-Extraposition and the Order of Adverbials in English.” *Linguistic Inquiry*: 1–50.

Richards, Norvin. 1999. “Featural Cyclicity and the Ordering of Multiple Specifiers.” *Current Studies in Linguistics Series* 32: 127–158.

Rohrbacher, Bernhard Wolfgang. 1994. “The Germanic Languages and the Full Paradigm.” PhD diss., University of Massachusetts.

Rohrbacher, Bernhard Wolfgang. 1999. *Morphology-Driven Syntax: A Theory of V to I Raising and Pro-Drop*. Amsterdam: John Benjamins.

Tvica, Seid. 2017. “Agreement and Verb Movement: The Rich Agreement Hypothesis from a Typological Perspective.” PhD diss., University of Amsterdam.

Vázquez Álvarez, Juan Jesús. 2011. “A Grammar of Chol, a Mayan Language.” PhD diss. University of Texas, Austin.

Part III. Explorations in Language Use across Modalities

Nonword Repetition in Czech-English Bilingual Children. Parallels with Developmental Language Disorder

Luca Cilibrasi [a] . D W H L Q D + ᵇ D VDQQDR Y$ £D @ E Ý W Ð % U D E

Faculty of Arts, Charles University, Prague, Czech Republic

[a]Luca.cilibrasi@ff.cuni.cz; [b]katkahasolova@gmail.com; [c]brabcova@gmail.com

Abstract: Bilingual children may display language difficulties that are, at times, misunderstood as language impairment. Previous work has suggested that nonword repetition tasks may be used to disentangle difficulties related to bilingualism from difficulties related to developmental language disorder (DLD), since the patterns observed in these two groups may be different: specifically, DLD children were shown to be more prone to errors related to nonword length and to be more sensitive to the presence of phonological clusters.

This study offers some preliminary data on a new sample of children, Czech-English sequential bilinguals, and it compares their performance to that of English DLD children (both groups were assessed with an English task). Contrary to previous studies, in this sample, groups did not differ in performance: both bilinguals and children with DLD in fact displayed a main effect of length (long nonwords were repeated less accurately), and a main effect of cluster (nonwords containing clusters were repeated less accurately). These findings suggest that the use of nonword repetition to disentangle bilinguals from children with DLD shall be exerted with caution and may not extend to all language pairs.

Keywords: bilingualism; nonwords; phonology; developmental language disorder

1. Theoretical Background

1.1 Nonword Repetition as a Clinical Tool

Nonword repetition is one of the most reliable systems to assess language difficulties in children and nonword repetition tasks are usually part of the assessment batteries used to diagnose a language disorder in a child. Classic research shows that deficits in

nonword repetition tasks correspond to deficits in working memory (Gathercole et al. 1994) and/or phonological processing (Snowling, Chiat, and Hulme 1991). Since both working memory and phonological processing appear to be impaired in children with language disorders, nonword repetition tasks have been shown to have an excellent diagnostic accuracy, and they are adopted across a wide variety of languages, such as English, German, Dutch, Italian, French, Russian, Czech, Vietnamese (Chiat 2015).

As Gathercole et al. (1994) explain, children, with or without disorder, are sensitive to the length of the nonword they are asked to repeat, and longer nonwords cause more problems and tend to be repeated less accurately. This finding is described as a working memory effect (Baddeley, 2003): When we listen to speech, the sequence of sounds perceived is rehearsed for a very short time in our mind, and then it is rapidly lost from memory. If the sequence of sounds has a meaning, its perception may lead to the activation of specific representations in the lexicon. If, as it is the case with nonwords, it has no meaning, the only representational component that will be active is working memory, and specifically the so-called phonological loop. Nonword repetition offers thus primarily a measure of the phonological loop (Baddeley 2003). While it is true that increased nonword length represents a challenge for all children (Gathercole et al. 1994), several studies show that these effects are larger in children with a developmental language disorder (Chiat 2015). Thus, difficulties associated with increased nonword length are considered a useful parameter to differentiate typically developing (TD)-children from children with DLD.

As further research has shown, the number of segments or the length in milliseconds are not, however, the only parameters that predict difficulty in nonword repetition. Archibald and Gathercole (2006), for instance, showed that nonwords matched in length may still lead to different results when they differ in their phonological complexity. Nonwords that contain phonological clusters (sequences of at least two consonants) are repeated significantly less accurately than nonwords that do not contain phonological clusters, and, again, this is particularly evident in children who received a diagnosis of language impairment. Thus, deficits occurring with increased nonword length are regarded as working memory deficits, while deficits occurring in nonwords containing clusters of consonants are regarded as phonological deficits, and these deficits may be XVHG WR DVVHVV D GHYHORSPHQWDO ODQJXDJH GLV Archibald and Gathercole 2007).

1.2 Nonword Repetition in Bilingual Children

Studies on the performance of bilingual children in the repetition of nonwords have given very mixed results. In principle, nonword repetition shall not be heavily affected by the domain that tends to be more problematic for bilingual children, lexical access (Bialystok, Craik and Luk 2008), since nonword are by definition items that do not belong to the lexicon. However, several studies have shown that TD monolingual children

outperform TD bilingual children in nonword repetition accuracy, and TD bilinguals may match monolingual children with DLD in their overall score with nonwords (Chiat 2015). This overlap is problematic, as it reduces the applicability of nonword repetition as a diagnostic tool. Due to this pattern in overall performance, several researchers have turned to a more fine-grained analysis of the features of the nonwords as predictors of accuracy, rather than looking at overall performance, when comparing monolingual and bilingual children.

The parameters presented in the previous section (length and phonological complexity) are important and may offer a tool to disentangle the difficulties observed in bilinguals and those observed in children with a language impairment. A study by Thordardottir and Brandeker (2013) is particularly relevant to explain this concept. In this study, the authors investigated nonword repetition scores in a large sample of French-English bilingual children and compared their performance to that of French monolingual children with a diagnosis of language impairment. In this experiment, the authors investigated how different nonword properties may affect performance, given that both bilinguals and DLD children may perform, in terms of overall accuracy, below the monolingual norm. The study showed that while DLD children were highly sensitive to nonword length, bilingual children were not affected by this property of the nonwords. Longer nonwords were problematic for children with DLD, but were not problematic for bilingual children. This difference may be regarded as a tool for language assessment. As the authors suggested, while the performance in terms of overall score may be lower in both bilinguals and children with DLD, differences of this kind may be used to understand whether a bilingual child has a disorder.

Another study that is promising in this regard is that of Dos Santos and Ferré (2018). In this experiment, the authors compared monolingual and bilingual children, with and without language impairment, in their nonword repetition scores. Children were assessed with a French nonword repetition task, and the bilinguals' sample consisted of children speaking French as L2 and a variety of languages as L1. Interestingly, the study in first instance displays a discrepancy with Thordardottir and Brandeker (2013): the length effect was observed not only in the DLD groups (both monolingual and bilingual), but also in the bilingual TD group. The only group that did not show a length effect was the monolingual TD group. As a second step, the authors compared the roles of phonological clusters (specifically comparing nonwords with one vs two clusters) in predicting performance. Their results show that this property was able to differentiate bilingual TD children from bilingual children with DLD. In this sample, both monolingual and bilingual DLD children displayed a sharp drop in performance when two clusters were present. On the contrary, the TD groups (both monolingual and bilingual) did not show a significant drop in performance. According to the authors, thus, phonological complexity may be used to disentangle difficulties related to bilingualism from difficulties related to a disorder: while overall scores may be lower in

both bilinguals and DLD children, only DLD children may show significant effects related to clusters.

This complex and mixed bundle of data offers ideas for further research and contributes to the quest for parameters that could help separate difficulties related to bilingualism from difficulties related to DLD. The final aim is finding a reliable system to assess DLD in bilinguals, with the rationale that if TD bilinguals and DLD children display qualitative differences in nonword repetition, when bilinguals show traits of DLD children in their performance one may suspect that they have a disorder. This concept was well summarized by Chiat (2015), in a thorough review of the potential role of nonwords in the assessment of language disorders in bilinguals. As she states (Chiat 2015, 14): "If bilingual children are less affected by length and/or syllable complexity than those with language impairment, these factors may help with clinical diagnosis in bilingual children when their overall scores fall below those of monolingual children."

As Chiat (2015) further explains, despite the promise of this idea, evidence is currently very limited, and further studies contributing to this research agenda are warranted. The current study contributes to this agenda with data from a previously unexplored group, Czech-English sequential bilingual children.

2. Methods

2.1 Ethical Concerns

This study is part of the Primus project "Core syntax in bilingual children with varying levels of input" (www.csbc.ff.cuni.cz), directed by the first author of this article. The project was reviewed by Charles University Ethics Committee, and it received ethical approval. This study contains testing involving human participants. All procedures performed involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments. Informed consent was obtained from at least one parent for all individual participants involved in this study.

2.2 Participants

The study involved two groups of children: one group of sequential bilingual children (Czech-English), and one group of monolingual English children with developmental language disorder (DLD).

7KH ELOLQJXDO JURXS FRQVLVWHG RI FKLOGUHQ tional schools in Prague. Participants were all bilingual speakers of English and Czech, having Czech as L1, and they have started acquiring English from the ages of 1 to 4 and are thus classified as early sequential bilinguals (Meisel 2009; Tsimpli 2014). The data for WKLV JURXS ZHUH FROOHFWHG E\ %UDEFRYi DV SDUW %UDEFRYi $ TXHVWLRQQDLUH GLVWULEXWHG WR W DJH RI RQVHW WR (QJOLVK DQG WKXV SURYLGHG WKH

tested a larger sample of children, but we decided to exclude simultaneous bilinguals because they were only 11, and the subsample was thus too small for us to reach meaningful conclusions. A parental consent for the testing was also provided from each participant. All children were Czech-English bilinguals, using mainly English in their education and both English and Czech in their everyday life (as assessed with a questionnaire).

The DLD group consisted of 18 children, aged 5 to 14, whose data was provided by the Department of Clinical Language Sciences at Reading University. These children attended the clinic because they were experiencing language difficulties, and they received a diagnosis of language impairment because they performed at least one standard deviation below the mean in at least one other language test. The tests used were the Test IRU WKH 5HFHSWLRQ RI *UDPPDU(%LVKRS WK)XQGDPHQWDOV(6HPHO :LLJ DQG 6HFRUG DQG 1989). Children in this sample were assessed with a battery of tests, and since these assessments were completed for clinical reasons and not for research reasons, each child did not necessarily complete the whole battery. Despite these differences, all children were assessed with the Children's test of Nonword Repetition (CNRep, Gathercole et al. 1994), ZKLFK LV DOVR WKH QRQZRUG UHSHWLWLRQ WDVN W background assessment of bilingual children, and at least one of the other language tests.

2.3 Nonword Test

The CNRep is one of the most widely used nonword repetition task in the United Kingdom (Archibald and Gathercole 2006). The task is comprised of 40 items of different length (10 two-syllable nonwords, 10 three-syllable nonwords, 10 four-syllable nonwords, 10 five-syllable nonwords). Example of nonwords used in this task are: "glastow" (2 syllables), "dopelate" (3 syllables), "woogalamic" (4 syllables), "sepretennial" (5 syllables). Normative data suggests that children of all ages tend to perform more poorly with nonwords of increasing length (Gathercole et al. 1994).

Our approach to the data follows Cilibrasi et al. (2018) and Archibald and Gathercole (2006), as it compares performance in nonwords with clusters against nonwords without clusters. Following Cilibrasi et al. (2018), we focus on noninitial clusters in medium and long nonwords (4 and 5 syllable), since initial clusters in short nonwords are shown to behave idiosyncratically and are a less reliable measure of phonological processing. Noninitial clusters are a more reliable measure of phonological processing, and in CNRep these clusters only appear in 4 and 5 syllable nonwords. Table 1 summarizes how this approach compares to the previous studies mentioned in the introduction. The three studies led to rather different results, but since both languages and methods differ, we may not reliably suggest what is causing these different outcomes. This summary may serve for further investigations in which a smaller number of conditions could be changed, and it may be possible in the future to establish whether differences between these three studies are related to methodological differences or to language differences (or both).

	Thordardottir and Brandeker (2013)	Dos Santos and Ferré (2018)	Cilibrasi et al. (2022)
Language	French	French	English
Target group	French-English bilinguals	French L2 speakers with and without DLD; mix of L1s	Czech-English bilinguals
Control group	French monolinguals with DLD	French monolinguals, with and without DLD	English monolinguals with DLD
Task characteristics	French nonwords of varying length (clusters not controlled)	French nonwords of varying length, with 0, 1 or 2 clusters	English nonwords of 4 or 5 syllables, with or without a cluster
Results	/ H Q J W K H u H F W L Q E R W K DLD and TD children	/ H Q J W K H u H F W L Q E R W K TD children. & O X V W H U H u H F W R O \ L Q DLD	Both groups show cluster and length H u H F W V

Table 1. Comparison of the current study with Thordardottir and Brandeker (2013) and Dos Santos and Ferré (2018).

2.4 Scoring

Bilingual children were asked to perform this task in their L2. Phonologies of L1 andL2 are shown to interact in bilingual speakers, and the phonology of L1 may be dominant over the other phonology, affecting the realization of phonemes in L2 production (Paradis 2001; Babatsouli and Ingram 2015). As such, scoring in this sample consisted in measuring whether the child was attempting at the correct target, rather than measuring whether the child was articulating with a standard British accent the phonemes they heard in the nonword. CNRep has a binary type of scoring: correct vs incorrect. In this study, deviations from the target were counted as mistakes, but not mispronunciations. For instance, the reduction of a cluster was counted as an incorrect repetition (i.e.“sepetennial” in place of “sepretennial”). In the meantime, the use of D W U L O O U W \ S L F D O R I &] H F K L Q W K H S O D F H R I D of British English, or similar phonemic variations, were not considered errors.

2.5 Hypothesis

Based on the previous work of Thordardottir and Brandeker (2013) and Dos Santos and Ferré (2018), our hypothesis is that bilingual children’s performance will be less dependent on phonological complexity and length than DLD children’s performance. Our prediction is that DLD children will be showing large length and cluster effects, while bilinguals will be showing small or nonsignificant effects for the same variables.

3. Results

Accuracy was measured as a binary variable (accurate vs inaccurate), and for each participant it was then coded as proportion of correct responses in each of the 4 given conditions (4-syllable with cluster, 4-syllable without cluster, 5-syllable with cluster, 5-syllable without cluster). Thus, for each condition, and each participant, accuracy was a value between 0 and 1. Descriptive statistics for both groups are presented in Table 2.

	Four syllables	Five syllables	Without cluster	With cluster
Bilinguals	0.79 (0.18)	0.69 (0.19)	0.78 (0.18)	0.7 (0.19)
DLD	0.53 (0.34)	0.41 (0.32)	0.53 (0.31)	0.41 (0.34)

Table 2. Descriptive statistics. Mean values for the variables investigated (standard deviation in parenthesis).

The full dataset was analysed with linear mixed models (Bates 2005). The chosen model included accuracy as the dependant variable, and length, cluster, and group as fixed effects (main predictors). The fixed effect of length consisted of two categories, 4 and 5 syllable nonwords (as noninitial clusters are only present in 4 and 5 syllable nonwords in the CNRep test). The fixed effect of cluster also consisted of two categories, presence vs absence of a non-initial cluster. Both length and cluster were re-leveled using the sliding contrast function (MASS library). The random effect of item could not be included because, for the clinical group, the data we had available were already summarised by condition (we did not have access to performance on each specific item), so the only random effect included was participant. The random structure was chosen comparing several possible random structures (using anova) and picking the model with the smallest AIC (see Baayen et al. 2008 for a description of this procedure). For a full list of the models compared see the Appendix. The chosen model was thus:

```
M2 <- lmer(acc ~ length*cluster*group + (length|part), na.action=na.omit)
```

The full model revealed a main effect of length, a main effect of cluster, and a main effect of group (Table 3). None of the interactions reached significance. Based on the signs of the estimates, these results indicate that long nonwords were repeated less accurately than shorter nonwords, nonwords without clusters were repeated more accurately than nonwords with clusters, and bilingual children outperformed in overall score the children with DLD. The lack of significant interactions indicate that the patterns observed are similar in the two groups: even though overall proficiency is lower in DLD children, both groups are similarly sensitive to length and cluster effects.

	Estimate	Std Err	T value	P value
Length	-0.11	0.02	-3.79	<.001 *
Cluster	-0.10	0.02	-3.88	<.001 *
Group	0.26	0.06	3.96	<.001 *
Length:Cluster	-0.05	0.05	-0.94	0.4
Length:Group	0.01	0.05	0.21	0.8
Cluster:Group	0.04	0.05	0.73	0.4
Length:Cluster:Group	0.13	0.11	1.26	0.2

Table 3. Fixed effects from the full model

To further explore the dataset, the two groups were also analysed separately, and age was added in each separate model. These models included accuracy as the dependant variable, length, cluster, and age as fixed effects (main predictors), and participant as a random effect. Age was centred, following directions of Kraemer and Blasey (2004).

The next section presents results obtained from these models, separately for the bilingual group and the DLD group.

For the Sequential Bilingual group, the model showed a significant main effect of length, t(63)= -2.81, p = 0.006, and a significant main effect of cluster, t(63) = -2.34, p = 0.02. No other main effect or interaction reached significance in the analysis. The two main effects are represented visually in Figures 1 and 2.

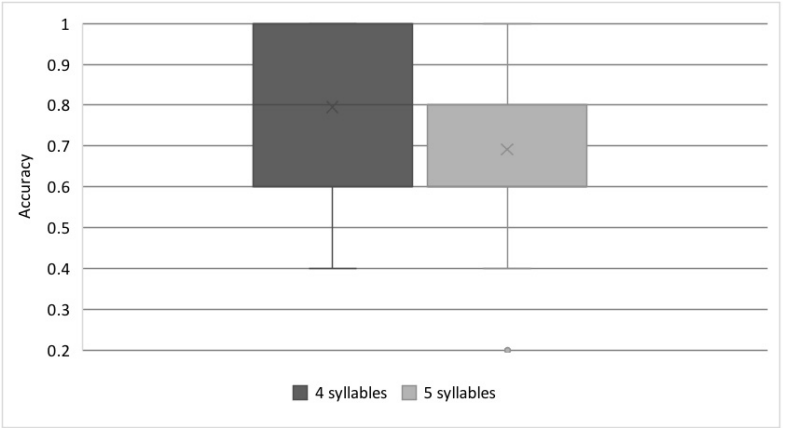


Fig 1: Proportion of correct repetitions in nonwords with 4 and 5 syllables in Sequential Bilinguals. This figure represents the main effect of length. Accuracy was obtained dividing the number of correct responses by the number of items repeated. Overall, nonwords with 4 syllables were repeated more accurately than nonwords with 5 syllables.

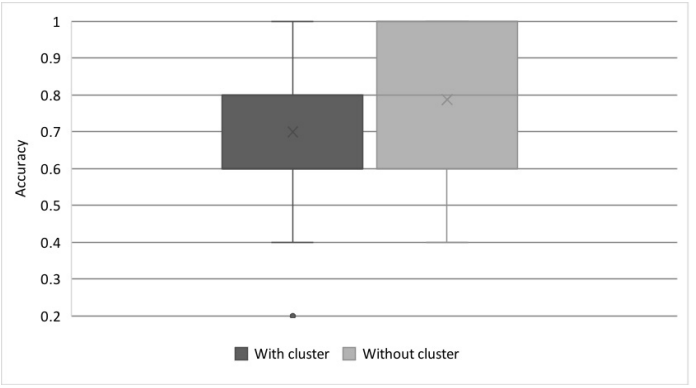


Fig 2: Proportion of correct repetitions in nonwords with either the presence or the absence of a word-initial cluster in Sequential Bilinguals. This figure represents the main effect of cluster. Accuracy was obtained dividing the number of correct responses by the number of items repeated. Overall, nonwords containing a cluster were repeated less accurately than nonwords which did not contain a cluster.

The second part of the analysis focused on the DLD participants. The model showed a significant main effect of length, $t(48)= -2.7$, $p = 0.009$, a significant main effect of cluster $t(48)= -2.96$, $p = 0.004$, and a marginal main effect of age $t(16)=2.02$, $p = 0,059$. No other main effect or interaction reached significance in the analysis. The two significant main effects are presented visually in Figures 3 and 4. The marginal main effect of age is not surprising in this sample, since the age range was considerably larger in the DLD group than it was in the bilingual group. The general tendency, as expected, was towards an improvement of performance with growing age.

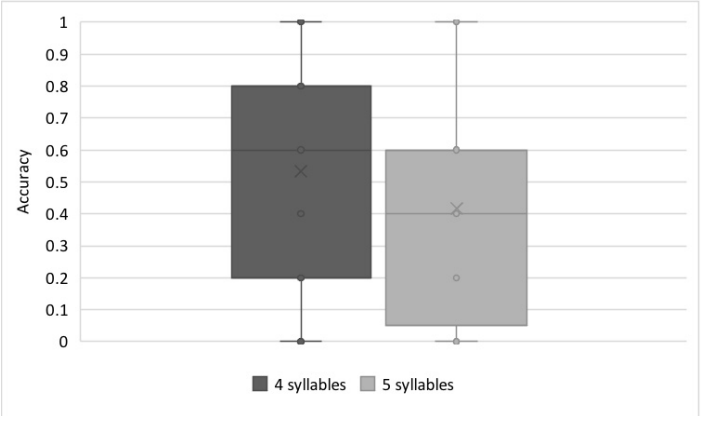


Fig 3: Proportion of correct repetitions in nonwords with 4 and 5 syllables in DLD children. This figure represents the main effect of length. Accuracy was obtained dividing

the number of correct responses by the number of items repeated. Overall, nonwords with 4 syllables were repeated more accurately than nonwords with 5 syllables.

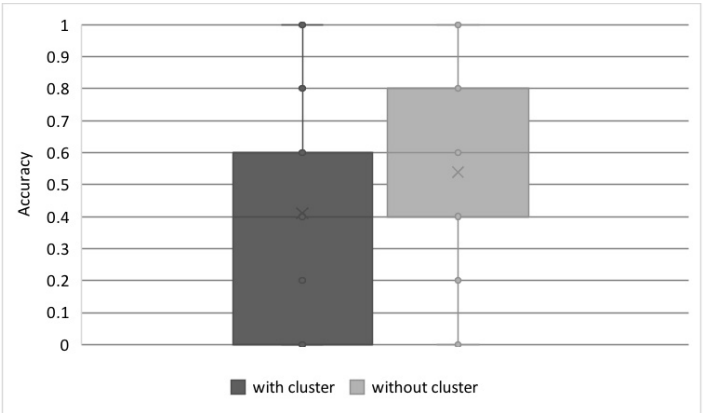


Fig 4: Proportion of correct repetitions in nonwords with either the presence or the absence of a coninitial cluster in DLD children. This figure represents the main effect of cluster. Accuracy was obtained dividing the number of correct responses by the number of items repeated. Overall, nonwords containing a cluster were repeated less accurately than nonwords which did not contain a cluster.

4. Discussion

This study investigated the roles of length and phonological complexity in predicting nonword repetition accuracy in a group of Czech-English sequential bilingual children, and it compared their performance to the performance of a group of monolingual English children with developmental language disorder (DLD). The results showed that children (in both groups) were sensitive to nonword length and nonword phonological complexity, with longer nonwords being repeated less accurately, and nonwords containing clusters being repeated less accurately. Bilingual children outperformed DLD children in overall score. Additionally, the DLD group displayed a marginal main effect of age, probably due to the larger age variability in that sample, with performance improving together with growth in age.

These findings contribute to the ongoing discussion on the potential role of nonwords in assessing DLD in bilingual children (Chiat 2015). In previous work, Thordardottir and Brandeker (2013) have shown that French-English bilingual children are less prone to length effects in nonwords than French monolingual children with a diagnosis of language impairment. This finding led to the proposal that length effects may be a good candidate for the recognition of qualitative differences in the performance of bilinguals when compared to children with DLD: the proposal of Thordardottir and Brandeker (2013)

is that, even though both groups may perform below monolingual norms, bilinguals may be differentiated because they are not sensitive to length effects. While a monolingual displays a length effect, one may suggest that this could be a flag of a potential language difficulty (a DLD symptom).

Our findings go against these claims and suggest that this approach may potentially be problematic. In our sample, both groups showed a main effect of length.

A second relevant comparison may be made with the study of Dos Santos and Ferré (2018). In their study, the authors found that phonological complexity was efficient in disentangling TD bilinguals from DLD monolinguals. In their sample, while monolinguals with DLD showed a sharp decline in performance when presented with nonwords that contain two phonological clusters, no significant decline in performance was observed in the bilingual sample. Importantly, a significant decline was instead observed in a group of bilingual children with DLD: this pattern was taken to justify the use of phonological complexity as a parameter to disentangle difficulties related to bilingualism from difficulties related to DLD. As a consequence, this discrepancy leads to the suggestion that phonological complexity may be used as a system to diagnose DLD in bilingual samples.

Also in this case, our findings are not consistent with previous work, and suggest thus that phonological complexity may not necessarily succeed in disentangling bilinguals and DLD children. Contrary to Dos Santos and Ferré, we did find that both groups, DLD and bilinguals, behaved in the same way with nonwords of varying phonological complexity: In both cases, we observed a main effect of cluster, with performance declining significantly when nonwords contained a cluster.

In summary, in the current study, bilinguals and DLD children were shown to be sensitive in nearly identical ways to the two parameters manipulated (length and phonological complexity). Thus, none of these two variables would have been efficient in separating these two groups, if deemed necessary.

The consequence is that (if such results were replicated) nonword length and nonword phonological complexity may not be safely used to recognize a language impairment in a bilingual child, at least in those with Czech L1 assessed in English L2. There is nothing in our bilingual sample that could make us suspect a language impairment in these children, and it appears rather that nonword length and nonword phonological complexity are affecting bilinguals' performance similarly to what is observed in DLD children. Following the reflections of Chiat (2015), we stress that further data is necessary to fully understand the extent to which nonword length and phonological complexity may be used to disentangle bilinguals from children with DLD, and ultimately whether some nonword properties may be used to assess DLD in bilingual children. Our study is smaller in size than the studies completed by Thordardottir and Brandeker (2013) and Dos Santos and Ferré (2018), and it relies on a different pair of languages.

Further work with larger samples and with a wider variety of languages may offer a clearer picture as to whether there are variables that can be used to disentangle between these two groups, and thus help in the assessment of DLD in bilinguals.

Funding Acknowledgement

This project was funded by the Primus Grant, Charles University, award 19/HUM/14.

Appendix

Models compared:

```
M1 <- lmer(acc ~ length*cluster*group + (1|part), na.action=na.omit)
M2 <- lmer(acc ~ length*cluster*group + (length|part), na.action=na.omit)
M3 <- lmer(acc ~ length*cluster*group + (cluster|part), na.action=na.omit)
M4 <- lmer(acc ~ length*cluster*group + (group|part), na.action=na.omit)
M5 <-  lmer(acc ~ length*cluster*group + (length*cluster|part), na.action=na.omit)
M6 <-  lmer(acc ~ length*cluster*group + (length*group|part), na.action=na.omit)
M7 <-  lmer(acc ~ length*cluster*group + (cluster*group|part), na.action=na.omit)
```

Models 5, 6 and 7 did not converge.

List of Stimuli:

	4 syllable	5 syllable
With a noninitial cluster	Contramponist, Perplisteronk, Stopograttic, Empliforvent, Blonterstaping	Sepretennial, Detratapillic, Confrantually, Underbrantuand, Versatrationist
Without a noninitial cluster	Woogalamic Fenneriser Commecitate Loddernapish Penneriful	Defermication, Reutterpation, Altupatory, Pristoractional, Voltularity

Works Cited

Archibald, Lisa M., and Susan E. Gathercole. 2006. “Nonword Repetition: A Comparison of Tests.” *Journal of Speech, Language, and Hearing Research: JSLHR* 49 (5): í

Archibald, Lisa M., and Susan E. Gathercole. 2007. “Nonword Repetition in Specific Language Impairment: More Than a Phonological Short-Term Memory Deficit.” *Psychonomic Bulletin & Review* 14 í

Baayen, Rolf Harald, Davidson, Douglas J., and Douglas M. Bates. 2008. “Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items.” *Journal of Memory and Language* 59: 390–412. Doi: 10.1016/J.Jml.2007.12.005

Babatsouli, Elena, and David Ingram. 2015. “What Bilingualism Tells Us about Phonological Acquisition.” In *Routledge Handbook of Communication Disorders*, edited E\ 5XWK +XQWOH\ %DKU DQG (ODLQH 5 6LOOLPDQ

Baddeley, Alan. 2003. “Working Memory and Language: An Overview.” *Journal of Communication Disorders* í

Bates, Douglas. 2005. “Fitting Linear Mixed Models in R.” *R News* 5 í

%U DE EFRY i $O åE WD ³(IIHFW RI 7LPH RI ([SRVXUH Relative Clauses in Primary School Children.” MA Thesis, Charles University in Prague.

Bishop, Dorothy V. 1989. Test for The Reception of Grammar (TROG). London: Medical Research Council.

Bialystok, Ellen, Fergus Craik, and Gigi Luk. 2008. “Cognitive Control and Lexical Access in Younger and Older Bilinguals.” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34 (4): 859.

Chiat, Shula. 2015. “Non-Word Repetition.” In *Assessing Multilingual Children*, edited E\ 6KDURQ $UPRQ /RWHP -RKQ + $ / GH -RQJ DQG Multilingual Matters.

Cilibrasi, Luca, Vesna Stojanovik, Tom Loucas, and Patricia Riddell. 2018. “The Role of Noninitial Clusters in The Children’s Test of Nonword Repetition: Evidence from Children with Language Impairment and Typically Developing Children.” *Dyslexia* 24 í

Dos Santos, Christophe, and Sandrine Ferré. 2018. “A Nonword Repetition Task to Assess Bilingual Children’s Phonology.” *Language Acquisition* 25 í

Gathercole, Susan E., Catherine S. Willis, Alan D. Baddeley, and Hazel Emslie. 1994. “The Children’s Test of Nonword Repetition: A Test of Phonological Working Memory.” *Memory* í

German, Diane J. 1989. Test of Word Finding (TWF). Dallas, TX: DLM.

Kraemer, Helena Chmura, and Christine M. Blasey. 2004. “Centring in Regression Analyses: A Strategy to Prevent Errors in Statistical Inference.” *International Journal of Methods in Psychiatric Research* 13 í

Meisel, Jürgen M. 2009. “Second Language Acquisition in Early Childhood.” Zeitschrift für Sprachwissenschaft í
Paradis, Johanne. 2001. “Do Bilingual Two-Year-Olds Have Separate Phonological Systems?” International Journal of Bilingualism 5 í
Semel, Eleanor M., Elisabeth H. Wiig, and Wayne Secord. 2004. CELF 4: Clinical Evaluation of Language Fundamentals 4 Screening Test. London: Pearson Clinical.
6 L O H R 5 R E H U W R D Q G ' D Q D 7 \ þ R Y i ³ &] H F N 5 H S , Task for Czech.” Studies in Applied Linguistics í
Snowling, Margaret, Shula Chiat, and Charles Hulme. 1991. “Words, Nonwords, and Phonological Processes: Some Comments on Gathercole, Willis, Emslie, and Baddeley.” Applied Psycholinguistics í
Thordardottir, Elin, and Myrto Brandeker. 2013. “The Effect of Bilingual Exposure Versus Language Impairment on Nonword Repetition and Sentence Imitation Scores.” Journal of Communication Disorders í
Tsimpli, Ianthi-Maria. 2014. “Early, Late or Very Late? Timing Acquisition and Bilingualism.” Linguistic Approaches to Bilingualism í

Eyebrow Raises as Facial Gestures: A Study Based on American Late Night Show Interviews

Volker Gast

Friedrich Schiller University, Jena, Germany
Palacký University, Olomouc, Czech Republic

volker.gast@uni-jena.de

Abstract: This study provides an exploratory analysis of eyebrow raises in a corpus of TV interviews. The central hypothesis is that brow raises, besides their prosodic, emotive and social functions, often have a metaphorical dimension, signalling 'openness'. This metaphorical function modulates aspects of meaning at various levels of linguistic interpretation. The multi-modal corpus data was biometrically annotated with facial landmarks. The audio signal was analysed in terms of pitch and intensity. The study combines high-resolution analyses of six speakers with analyses of the whole corpus. The hypothesis that brow raises metaphorically signal openness receives some support from the data, especially with respect to their occurrence in the context of epistemic modality and additivity. A statistical association with polar questions cannot be observed. The study also demonstrates that more data, observational as well as experimental, is needed to gain a more complete understanding of the function of facial gestures in spoken conversation.

Keywords: facial gestures; multimodal corpus; eyebrow raises; metaphor

1. Introduction

While there is a substantial body of literature on facial gestures from a psychological point of view, mostly with a focus on the expression and processing of emotions (e.g. Ekman 2003), the "semantics of facial expressions" (Wierzbicka 2000, 15) is more poorly understood. In their Integrated Message Model of Language, J. Bavelas, N. Chovil and colleagues have proposed treating Conversational Facial Gestures as a component of the gesture-speech ensemble (Bavelas and Chovil 1997, 2000, 2006; Bavelas and Gerwing 2007; Bavelas et al. 2014a,b; Bavelas and Chovil 2018). It is in this spirit that the present article proposes an analysis of eyebrow movement, drawing

on research on hand gestures in spoken language (McNeill 1992, Kendon 2004) as well as non-manuals in sign languages (e.g. Pfau and Quer 2010, Herrmann and Steinbach 2013 across languages, and Bross 2020, Pendzich 2020 on DGS).

The study addresses the following question:

(1) Under what circumstances do speakers raise their eyebrows?

Wierzbicka (2000), referring to earlier work by other authors, associates raised eyebrows with "attentional activity", "an effort to recall", an "act of empathy", "novelty", "surprise", "doubt", "incredulity", "disbelief", and "interest" (cf. Wierzbicka 2000, 164ff. for references). She proposes the following Natural Semantics Metalanguage representation:

(2) I know something now.
I want to know more (about this).
I'm thinking now.
(Wierzbicka 2000, 168)

An empirical study of eyebrow movement was carried out by Flecha-García (2006, 2010). Using the framework of Conversational Game Analysis (Carletta et al. 1997), she found that raised eyebrows can be observed in "Instruct" moves (though not in "Query" moves), and at the beginning of higher-level turn-constructional units ("Transactions"). Flecha-García (2006) also identifies correlations between eyebrow movement and pitch contour (see also Guaïtella et al. 2009 and earlier work cited there, and Kim et al. 2014).

The present approach is located in between the more general description of Wierzbicka (2000) on the one hand, and the more specific analysis of Flecha-García (2006, 2010), on the other. I intend to provide an abstract, general meaning of (specific types of) brow raises, deriving their particular communicative effects from a combination of that general meaning with properties of the context.

The central hypothesis of this study is that eyebrow movement is metaphorically associated with openness in some of its uses. Openness is regarded as a relative notion. It is defined in (3).

(3) A category A is more open than a category B iff the instances of B form a real subset of the instances of A.

I assume that facial gestures – like hand gesticulations – modulate the information conveyed through the acoustic signal. Brow raises may indicate a higher degree of openness in comparison to the default category. For instance, a phonetically encoded proposition comes with a specific amount of epistemic commitment. By using (default) indicative mood, making an epistemically unmodified statement, a speaker signals that

they fully commit to the information in question (e.g. The road is blocked). Epistemic commitment can be reduced by using a modal (The road may be blocked). The indicative sentence reduces the possible worlds to those where the road is blocked, whereas the modal may leaves all possibilities open (the road may or may not be blocked). As modulators of modality, eyebrow raises thus typically accompany expressions of uncertainty, rather than certainty.

I use an annotated multi-modal corpus of TV late night show interviews to address the question in (1). This corpus is described in Section 2. In Section 3, some prosodic, emotive and social functions of eyebrow movement are briefly addressed. Section 4 deals with the central topic of this study, the metaphorical function of brow raises. Section 5 contains the conclusions.

2. Corpus and Methods

The corpus consists of video material with interviews conducted in the Late Show with David Letterman between 1980 and 2014, collected from a fan page, where the videos can be accessed.[1] It comprises 160 video files (episodes) with an overall duration of ~ 160 hrs, with automatically generated subtitles. The files were processed as follows:

- The audio signal was extracted (with ffmpeg[2]) and intensity and pitch contours were measured with Praat (Boersma and Weenink 2018).
- Rectangles with faces were identified using the 'dlib'-package for Python[3].
- For rectangles with a size of at least 10,000 pixels, 3D facial landmark detection was applied (see Figure 1), using the 'face alignment' package for Python[4].
- The facial landmark coordinates were normalized to the width of the eyes, the height of the nose and the width of the nose.
- The position of the eyebrows was measured as the distance between the centroid of the ten facial landmarks of the eyebrows, and the centroid of the five facial landmarks of the nose.
- The sagittal and vertical head angles were determined (this was possible because the 'face_alignment' package delivers a three-dimensional model of the head).

1 https://donzblog.home.blog/

2 K W W S V Z Z Z u P S H J R U J

3 https://pypi.org/project/dlib/

4 https://github.com/1adrianb/face-alignment

The basic unit of observation was the frame. As the video data was processed at a rate of 30 frames per second, each second is associated with 30 (sequential) frames. Each frame is associated with five measurements: (i) intensity and (ii) pitch of the audio signal, (iii) height of the eyebrows, and (iv) sagittal as well as (v) vertical angle of the head. The resulting dataframe was reduced to those observations that were part of an uninterrupted sequence of at least 90 frames (3 seconds) showing the same speaker, filmed from the same angle. The audio-visual data was aligned with the subtitle units as represented in WKH VXEWLWOH IOHV 6XEWLWOH XQLWV KDYH D PH WRNHQV RQ DYHUDJH 1 $OO PHDVXUHPHQW over those units, determining mean values, maximum values and minimum values for eyebrow height and pitch. Sequences of frames of this type will be called 'segments'.

The multimodal corpus can be represented using the ANVIL software (Kipp 2001) as shown in Figure 2. As the top-right box in Figure 2 illustrates, segments are not (necessarily) meaningful units of speech (here, what the story was in fact Miss Wyman, corresponding to segment 0250 of episode 0188, from 00:11:32,100 to 00:11:35,130).

Given that the subtitle units (segments) are relatively short, I created an additional data frame that contains the measurements aggregated over each segment as well as the following segment. The reason is that some effects can be expected to show with a certain time lag. For example, focus particles and other scope-bearing elements may be associated with eyebrow activity later in the signal if we assume that it is not (necessarily) the particle itself, but the focus, that triggers eyebrow movement. The second data frame taking into account the potential temporal delay of eyebrow movement in relation to the verbal material will be called the 'lag data'.[5]

5 The data is available in the following OSF repository: https://osf.io/y5s2u/

As eyebrow movement varies considerably across speakers, I have combined analyses of the entire corpus with a more fine-grained approach, inspecting six individual speakers, who will be identified with their initials in the following: Eddie Murphy (EM), Teri Garr (TG), Art Donovan (AD), Buck Henry (BH), Quentin Crisp (QC) and Norman Macdonald (NM). David Letterman (DL) mostly figures as host, but there are also interviews conducted by Tom Snyder (TS).

3. Non-Metaphorical Functions of Eyebrow Raises

The hypothesis of brow raises being associated with openness applies to descriptive meaning only (see Sect. 4). Non-metaphorical occurrences of brow raises can be motivated by prosody (Sect. 3.1), emotions (Sect. 3.2) and the negotiation of social relations during conversation (Sect. 3.3).

3.1 The Prosodic Function of Brow Raises

Like hand gestures (esp. beats), facial gestures may convey prominence. Eyebrow raises are often used as "emphasizers" (Ekman 1979), "which [occur] synchronously with a stressed (e.g., prosodically marked) word" (Bavelas et al. 2014a, 124). In (4), the brow raise is precisely aligned with the focal accent. (Here and in the following, brow raises are indicated by a line over the relevant material, like this. Prosody is approximately annotated with small caps indicating prominent syllables and diacritics showing the type of tone in terms of Wells 2006, e.g. FÀLL, RÍSE. The examples can be inspected by clicking on the annotated material [in blue]).

(4) DL: ... you've handed me a little gift here a little pamphlet it says mental game and this is a little uh brochure, a little pamphlet what is it it's from ...
TG: Ja %‰+.
DL: ... from Japan, yeah what doe- what does this mean …

(5) is a textbook example of topicalization in the sense of Birner and Ward (1998). Brick work has been previously mentioned, and it stands in a paradigmatic contrast to RWKHU WSHV RI PDQXDO ODERXU SRWHQWLDOO\ ILO x type of labor].

(5) NM: I did a lot of manual labor, did all manual labor.
TS: Really?
NM: Yeah.
DL: Like brick work?
NM: % ‹ § ¦„work I } ¦ }, yeah!

As mentioned above, speakers differ considerably in terms of their individual eyebrow movement patterns. Globally, there is a clear positive correlation between the maximum values of eyebrow height and the maximum values of pitch per segment ($p<0.001$, according to a linear regression model with a random intercept for 'speaker'[6]). However, a linear model mapping maximum pitch to maximum eyebrow height is probably not an appropriate way of capturing the relationship between the two variables. While one speaker shows a rather linear pattern (AD, cf. the left plot in Figure 3), other speakers exhibit a positive correlation only in the lower areas of pitch (e.g. TG, cf. the right plot in Figure 3).

Figure 3: Relationship between maximum pitch (x-axis) and maximum eyebrow height (y-axis) per segment for two speakers, with a ~ line (in blue); all values are centered and scaled per speaker.[7]

Obviously, the prosodic function of brow raises is not always easy to keep apart from other functions. For instance, if a speaker stresses a modal (The road "be blocked), an accompanying brow raise could reflect either the pitch accent as such (and the focus marked by it), or the modal qualification (or both).

In Section 4, the prosodic function of brow raises will be taken into account by including 'pitch' as a control variable in the regression models predicting eyebrow height on the basis of variables reflecting semantic properties of the context (modality, sentence type, additive operators).

---

6 The model was fitted with the function lmer() (Bates et al. 2015).

7 The plots were generated with the R-package 'ggplot2' (Wickham 2016).

3.2 Emotive Functions of Eyebrow Raises

The archetypical trigger of raised eyebrows, it is often assumed, is the mental state of attention (e.g. Darwin 1872, Eibl-Eibesfeldt 1972). Attention, in turn, is often associated with the emotions of surprise and astonishment. An example where brow raise is most probably motivated by surprise is given in (6). The speaker learns during the interview that Sophia Loren will be attending an event that he will also be attending. The (surprised) questions Is she gonna be there? and Sophia Loren? are accompanied by brow raises. The "attentional activity" (Wierzbicka 2010) is also reflected in forward lean during Sophia Loren?.

(6) DL: Sophia Loren is gonna be there?
EM: Is she gonna be there?
DL: Ritchie?
Ritchie: Yes.
EM: Sophia Loren?

In (7), the speaker expresses astonishment and admiration for the acting skills of a colleague. As can be seen when the material is inspected, the main gesture is probably a lid raise (indicated by the dotted line on top of very very), with the speaker opening his eyes wide.

(7) EM: ... and I'm in a movie with an actor called Michael Wincott,
who's a very very good actor,
and Michael Rapaport, directed by Thomas Carter,
I can't wait till you see this one, you know ...

Emotions such as surprise and astonishment cannot easily be operationalized in a corpus based study. One may think of specific linguistic markers such as really? as easily indicators of surprise. While the corpus contains 95 occurrences of this adverb, they are often uttered by the host (D. Letterman) in reaction to something a guest has said, and the host is typically not in the picture. I can therefore not offer any quantitative data concerning a correlation between surprise and brow raises on the basis of my corpus. In order to study the position of eyebrows relative to specific emotions, controlled experiments will be needed (cf. also Sect. 5).

3.3 Social Functions of Eyebrow Raises

As a resource for 'rapport management', used to negotiate social relations between interlocutors, raised eyebrows may signal a positive attitude towards the conversational contribution made by the interlocutor. Consider the dialogue in (12). When giving a negative answer (Do you own a dog? – No.), the speaker does not raise her eyebrows. She

does so when she provides an affirmative answer though (Have you ever owned a dog? Yes!). Her response to the host's statement Dogs are great, I think – Hm hm, yeah, that's right – is also accompanied by a noticeable brow raise.

(8) DL: ... and, uhm, a guy named Jim Buck will be here. Jim makes his living walking dogs in the city of New York. Do you have a dog?
TG: No.
DL: Have you ever owned a dog?
TG: Y ¢ Œ!
DL: Dogs are great, I think.
TG: Hm-hm, ' ¢ z •, that's ‹ ¦.€ • •

A positive stance cannot only be expressed towards the conversational contribution but also towards the interlocutor themself. In this case, brow raises can establish or enhance rapport in a more general way. In (9), the speaker raises her eyebrows for a pronounced gesture, reflecting the rapport established in the conversation. Such uses often have a 'conspirational' effect, pointing to shared private knowledge.

(9) DL: ... I was surprised to learn this afternoon that you attended an event here in New York City that we've covered about four three or four times ...
TG: Small world, isn't it?
That we { « • were ¦ ‡terested in this
DL: Yeah, the inventors' exhibition ...

Rapport management can probably be studied systematically on the basis of observational data. However, it would require filming both interlocutors, ideally from various angles, and such analyses would probably require comprehensive manual annotations. The social function of brow raises can therefore not be taken into consideration in the present (quantitative) study. I assume, however, that brow raises reflecting rapport as in (9) are in principle independent of the variables investigated in Sect. 4 and therefore do not have to be controlled for.

4. Eyebrow Raises Modulating Descriptive Meaning

The argument made in this section was already laid out above: I argue that brow raises often – not always – metaphorically convey openness as defined in (3). They typically modulate an existing element of sentence semantics. This modulation may happen at various levels of (non-emotive and non-social) interpretation. My approach has been inspired by F. Bross's and D. Hole's work on German Sign Language (DGS). Bross and Hole (2017) argue for an isomorphism between the position of non-manual articulators in the human body and their scope (the "Bodily Mapping Hypothesis"). The hypothesis

pursued in the present study is different from the Bodily Mapping Hypothesis, however. I assume that eyebrow movement may interact with different levels of meaning, and that its specific effect will be different in each case.

To structure the presentation of the data I use the metafunctions distinguished in Systemic Functional Grammar (e.g. Halliday 1994) (rather than the generative framework used by Bross and Hole 2017). At the highest level of the taxonomy Halliday (1994) distinguishes three functions, (i) the ideational function of language (Sect. 4.1), (ii) the interpersonal function (Sect. 4.2), and (iii) the textual function (Sect. 4.3). Since brow raises seem to operate primarily at the interpersonal and textual levels, Sect. 4.1 is considerable shorter than the other two sections. As will be seen, it is doubtful whether brow raises have a function at an ideational level at all.

4.1 The Ideational Level

Hand gestures are often used metaphorically, e.g. temporal development is accompanied by horizontal movement of a hand (see for instance Cienki and Müller 2008). Facial gestures do not commonly seem to be used in this way. We may expect them to accompany expressions denoting the higher region or end of a scale, e.g. indicating large quantities or – in accordance with the main hypothesis of the present study – openness of some container. However, I have not found any striking examples of this type in my corpus. A candidate for a scalar context was given in (7) above. The speaker (EM) emphasises the high quality of a referent's acting (a very, very good actor), and raises his eyebrows while uttering very. But then, as has been pointed out, the main gesture is a lid raise, which is probably motivated by the emotion of admiration.

In order to test whether the expression of a high degree (at a propositional level) is associated with brow raises I determined the distribution of eyebrow height relative to the (non-)occurrence of the scalar adverb very. The data of my corpus do not show any statistical association of very with high eyebrow positions (controlling for pitch). For none of the three ways of aggregating over the data per segment – mean values, maximum values or minimum values – can a significant correlation be observed.[8] Figure 4 shows the distribution of mean values for segments with and without very, centered and scaled per speaker. The violin plots show that the two distributions – segments with and without very – are virtually indistinguishable.

8 The dataset contains 88 segments with very and 5,113 without very (after removing outliers). / L Q H D U U H J U H V V L R Q P R G H O V Z L W K µ S L W F K ¶ D Q G R µ Y H U \ ¶ D V G R Q R W U H Y H D O D Q \ V L J Q L ¿ F D Q W D V V R F L D W L R Q E H W Z H H Q W K H H \ H E U R Z K H L J K W (mean values, p P D [Y D O X H V P L Q Y D O X H V

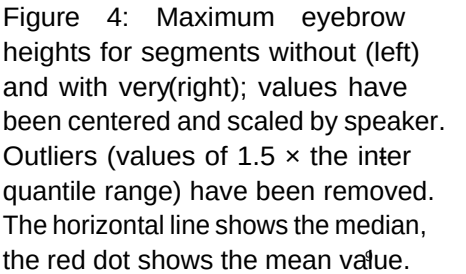


Figure 4: Maximum eyebrow heights for segments without (left) and with very(right); values have been centered and scaled by speaker. Outliers (values of 1.5 × the inter-quantile range) have been removed. The horizontal line shows the median, the red dot shows the mean value.[9]

4.2 The Interpersonal Level

The interpersonal function covers the domains of modality, mood/illocutionary force and polarity, according to Halliday (1994). Polarity will not play a central role in the following discussion, even though it may interact with other categories, e.g. insofar as uncertainty is often expressed using negation (e.g. I don't know). The discussion focuses on epistemic modality (Sect. 4.2.1) and sentence type (Sect. 4.2.2).

4.2.1 Epistemic Modality: Expression of Uncertainty

A common context for brow raises is the expression of uncertainty. A relevant example is given in (10), where uncertainty is made explicit (I'm not sure). The peak of the raise is here aligned with the negation marker not.

(10) DL: ... now let's talk about your association with Falcon Crest, you explained to me that you're not necessarily joining the cast.
BH: I'm actually not a member of the cast, uh, I joined it as a cameo performer for either three or four shows,
I'm ‡ « • Œ ° ‹ ~.
I actually haven't ever seen Falcon Crest.

Uncertainty or, in fact, ignorance, is systematically encoded in an emblem in which the eyebrows take part, the "facial shrug" (see for instance Bavelas et al. 2014a, Bavelas and Chovil 2018). The facial shrug at the same time often conveys indifference. An example is given in (11) (the dashed line over I guess so indicates the facial shrug).

9 The plots were generate with the R-package 'ggstatsplot' (Patil 2021).

(11) DL: Was it a success, New Year's Eve?
AD: [illegible] They all got drunk, uh, my wife ...

The distribution of eyebrow height in segments with and without markers of epistemic modality is shown in Figure 5. The category of 'epistemic modal' here comprises the modals may (n [illegible] D V Z H O O D mayb (n [illegible] D V Z H O O D maybe (n [illegible] D V Z H O O D perhaps (n [illegible] D V Z H O O D probably (n [illegible] 7 K H W K U H H S O R W V V K R Z W K H G L V W U L E X values and minimum values, from left to right and top to bottom.

Figure 5: Distribution of values for segments with and without an expression of epistemic modality, for mean, maximum and minimum values, from left to right and top to bottom; all values have been centered and scaled by speaker. Outliers (values of 1.5 × the interquantile range) have been removed. The horizontal line shows the median, the red dot shows the mean value.

Linear regression models with the occurrence of an epistemic modal and the pitch measurements as main effects and random intercepts for 'speaker' show a significant difference can only be observed for minimum values per segment (p E X W Q R mean values (p R U P D [L P X P Y D O X H V p 7 K H G D W D an epistemic modal and 5,099 segments without one (after removing outliers).

The results suggest that epistemic modality is not necessarily associated with individual, marked eyebrow raises, but rather with the absence of particularly low eyebrow positions. Note also that the plot showing the minimum values for pitch and eyebrow height at the bottom of Figure 5 actually suggests a bimodal distribution. I am unable to interpret these findings on the basis of my corpus data. There may be additional variables at play which my current dataset does not capture. I do assume, however, that the statistical association of epistemic modals with the absence of particularly low eyebrow positions is robust.

Expressions of epistemic modality satisfy the criterion of (relative) openness, in comparison to indicative modality, because the propositions (sets of worlds) that they denote form supersets of the worlds denoted by the latter category (see for instance Kratzer 1978 for a treatment of modality in terms of possible world semantics).

4.2.2 Sentence Type: Polar Questions

In Halliday's (1994) terms, the category of 'mood' covers illocutionary force, sentence type and polarity. This section focuses on sentence type. The most important sentence type in the context of brow raises is the one of 'polar question'. Brow raises have been claimed to be associated with polar questions in sign languages (e.g. Zeshan 2004, Cecchetto 2012) as well as in spoken language (e.g. Ekman 1979, Chovil 1991a, Chovil 1991b, Kim et al. 2014). Flecha-García (2006) found no evidence for brow raises as concomitants of questions, however. In fact, Kimmelmann et al. (2020) also found that raised eyebrows in Kazakh Sign Language primarily signal surprise, not illocutionary force.

The use of brow raises in the context of questions would be compatible with the hypothesis of this study, i.e. that brow raises may metaphorically signal openness. Questions have been analysed as denoting sets of propositions, more specifically, sets of possible answers (Hamblin 1973). The superset condition specified in (3) is therefore met, as the denotation of a question is a superset of the denotation of the corresponding declarative sentence. Note also that, while 'question' is primarily an illocutionary category, it is associated with a modal attitude of uncertainty or ignorance.

Before we look into polar questions specifically note that it is hard to separate the sentence type of 'polar question' from the emotion of 'surprise', as questions often come with surprise (specifically in the genre used for this study, TV interviews). For instance, example (6) above (Is she gonna be there? ... Sophia Loren?) is a raised question, and it is hard to tell whether the raised eyebrows modulate the illocutionary force 'question' or the emotion 'surprise', or both. (12) is an example where an eyebrow raise signals

a question without at the same time indicating surprise. The facial gesture is actually aligned with a pause and precedes the explicit question (Do you?).

(12) BH: ... but, but uh, I do what normal people do, when I'm uh, when I'm not doing whatever abnormal people do. Uhm, I go to films, I go to plays – just like you do.
<pause>
Do you?
DL: Yeah.

(13) is a similar example, insofar as the speaker raises his eyebrows before asking a question. In this case, however, there is probably surprise involved.

(13) DL: Uh, now, I understand uh from, uh, looking over some information this after-noon – documents, heretofore classified – uh, that you, you go for periods of of uh, insomnia – relentless insomnia ...
BH: <pause> Is • •• in the } «|uments?
DL: Yes, it is.

In order to test whether polar questions are associated with raised eyebrows, I identified clauses whose subject was immediately preceded by an auxiliary, and the auxiliary was not preceded by a wh-pronoun, as an approximate operationalization of polar questions. Such segments do not show higher eyebrow positions than other segments.[10] The distributions are shown in Figure 6.

The operationalization of polar question used in this study (subject preceded by auxiliary not preceded by wh-pronoun) may be a coarse-grained, but it is certainly not entirely off the mark. Given the complete absence of any kind of discernible asymmetry in the data (with p-values of 0.9, 0.57 and 0.34 for mean, max and min-values), it seems unlikely that polar questions should show any association with eyebrow position. Still, more precise measurements will be needed before the assumption that brow raises reflect polar questions can be discarded.

10 The annotated dataset contains 363 segments with a polar question and 5,211 segments without a polar question (after removing outliers). Linear regression models with 'polar question' DV PDDLQ HuHFW DQGUDQGRPVHDPHUHFWDWGDR QRW UHYHDO DQ\ VLJQLILFDQW presence of a polar question in a segment on eyebrow height (mean: p p PD[PLQ p

Figure 6: Distribution of aggregated eyebrow height values for segments with and without subject-auxiliary inversion (without preceding wh-pronoun), for mean, maximum and minimum values per segment, from left to right and top to bottom; all values have been centered and scaled by speaker. Outliers (values of 1.5 × the interquantile range) have been removed. The horizontal line shows the median, the red dot shows the mean value.

4.3 The Textual Function: Additivity

In their textual function, brow raises are typically forward-looking, and they can often be observed at the beginning of turns (Flecha-García 2006, Guaïtella et al. 2009). A typical context of this type is provided by the discourse marker well. A relevant example is given in (14). Note that well is also accompanied by a slight backward thrust of the head in this example. Ample context is provided for the reader to get a better understanding of the conversational embedding.

(14) DL: What was the motivation for you personally to take the gig as we say?
HB: An enormous amount of money. It was only like three or four days and I said, "Will you give me almost everything in the world?", and they said "Sure", so I did id.
DL: Why not, exactly!
HB: And it was fun. They were extremely nice people. They fed me; and they, uh, offered to drive me to the studio, an offer I didn't accept. I know better than that.
DL: Yeah. Now, why do you know better than that?
HB: W ¢ … .‡,~‹ I live in a‰ … › | ~ where nobody can • ¦ ‡ } me.
So I know that they can never get a car there. And two, people have disappeared on those trips

An example where a brow raise occurs at the beginning of a turn-internal episode, though not at the beginning of a turn, is given in (15).

(15) NM: ... the people that I worked with, man, they loved lifts and furniture.
« ‡ ~ time, one time, • • ¢ ‹ as interesting Œ • time, it's ‰ probably too ..‡ €
but ...

The projecting function of eyebrow raises can also be observed at a lower level, in lists. Consider (16), where the speaker provides a two-items list of examples as an answer to the question What constitutes a good model? On the first element (You could be handsome), he raises his eyebrows, along with a final fall tone. After a brief pause, he mentions a magnificent physique as a second quality. The brow raise on the first item seems to indicate that the list is not yet complete. A similar effect can be observed on the second item (magnificent physique), showing that the list is still not exhaustive.

(16) DL: And what constitutes being a good model?
QR: Well, I suppose you could have various things to offer.
You could be • œ ‡ } some
or you could have a magnificent physique ...

The type of list illustrated in (16) points to the more general function of 'conjunction' or 'additivity' as a context for brow raises. This hypothesis can be tested by determining the distribution of eyebrow heights relative to the conjunction and. A mixed linear model taking into account all occurrences of and (n_{and} = $n_{‡ ^ and}$ = D I W H U U H P R Y H outliers), with 'pitch' and 'and' as main effects, and 'speaker' as a random effect, shows a clear association of this conjunction with values for maximum eyebrow height within segments ($p<0.001$), though not for mean values (p = R U P L Q L P X P Y D O X H

Inspection of the data shows, however, that the raises are rarely, if ever, aligned with and itself; rather, it seems to be the forward-looking function and, hence, the often unit-initial position of brow raises that associates them with it. In any case, brow raises seem to be associated with the function of 'additivity' at a textual (syntagmatic) level.

The function of additivity also seems to play a role at an information structural level. Specifically, additive focus particles and similar expressions seem to be associated with raised eyebrows (see for instance Herrmann 2013 on the role of brow raises in three sign languages). If we pool the additive focus particles (scalar, n [illegible] (non-scalar, n [illegible] with either (n [illegible] in a linear mixed effects model with the measurements for 'pitch' and 'eyebrow height' as main effects, and 'speaker' as a random effect, shows a significant association between the occurrence of additive operators and mean as well as minimum eyebrow height (mean values p [illegible] ; p < 0.01), in the lag data (where measurements from the current segment as well as the next segment are taken into account). This correlation cannot be observed for maximum eyebrow height though (p [illegible]) (figure shows distributions.

An example of a brow raise in combination with a scalar particle (even) is given in (17).

(17) BH: Uh, [illegible] even at a [illegible] times to ex[illegible]
what I do when I'm not doing whatever it is I do that you know about.

An important question in the study of brow raises in the context of focus particles is to what extent this connection is sensitive to (esp. information structural) factors other than additivity. In particular, focus alternatives are often ordered, obligatorily so in the case of (scalar) operators such as even, optionally in the scope of other operators such as also (cf. Gast and van der Auwera 2011). Brow raises may thus also be triggered by modal attitudes coming with scalar orderings, e.g. of (un)likelihood. Once again, more fine-grained annotations and controlled experiments will be needed to address this question on an empirically solid basis.

I take it that the function of additivity instantiates openness as defined in (3), as it implies the broadening of a set of alternatives and the addition of propositions sharing a background to the Common Ground. In the case of turn- or episode-initial occurrences, the raises seem to be forward-projecting, scoping over the following discourse unit as it were.

Figure 7: Distributions of values for segments with and without an additive operator, for mean, maximum and minimum values, from left to right and top to bottom; all values have been centered and scaled by speaker. Outliers (values more than 1.5 × the interquantile range) have been removed. The horizontal line shows the median, the red dot shows the mean value.

5. Conclusions

The starting point of the present study was the hypothesis that in specific uses, brow raises are motivated metaphorically by signals of openness. This hypothesis does not imply that brow raises have no other functions, of course. Examples of prosodic, emotive and social functions were discussed in Section 3. While the corpus data presented in this article, as I hope, lends some plausibility to the 'openness hypothesis', it has certainly become clear that a sound empirical study of the conditions under which speakers move

their eyebrows requires more data – both more comprehensive and more richly annotated corpus data, and data from controlled experiments. Nevertheless, I feel that the following results obtained in this exploratory study seem to be more or less robust and can therefore be taken as a point of departure for future investigations:

- The expression of epistemic modality seems to correlate with eyebrow movement. While my data only show the minimum values per segment to be significantly higher in the presence of an epistemic modal (may, might) or adverb (maybe, perhaps, probably), not for mean or maximum values, that result merits further investigation.
- No significant association of brow raises with polar questions could be found. Before drawing any definite conclusions in this respect, more thorough investigations, based on more precisely annotated data, will be needed, however.
- Raised eyebrows seem to correlate with the textual or information structural function of 'additivity'.

In the text itself as well as in the summary given above, a few desiderata have been pointed out. The most important ones are:

- More multimodal corpus data is needed, ideally with cameras filming all interlocutors separately, to capture any interactional features.
- More fine-grained annotations are required, both automatic (higher resolution) and manual (capturing, for instance, emotions, social interaction and specific types of semantic or pragmatics contexts).
- Controlled experiments will be needed to answer specific questions relating to the movement of eyebrows.

As far as additional multi-modal corpus material is concerned, more data is needed simply because some of the contexts of interest (e.g. epistemic modals) are not particularly frequent, at least in the genre analysed for this study (TV interviews). Conversational games and similar elicitation tasks could be used to generate data with a number of occurrences of specific linguistic categories, with a certain amount of control (e.g. eliciting polar questions, modals and focus particles). Also, textual data will have to be aligned more precisely with the audio signal, using forced alignment. This would help us to get a better understanding of the timing of brow raises. More precise syntactic, semantic and ideally also pragmatic annotations would make it possible to address some more specific questions, e.g. where exactly brow raises are located relative to focus particles (e.g. on the particle itself, in its scope, or on the focus). As for experimental data, it will be instructive to differentiate systematically between specific triggers of eyebrow movement, such as questions and surprise (see also Kimmelmann et al. 2020 on Kazakh Sign Language).

The main challenges in the study of facial gestures at the current state of knowledge are thus methodological ones; but the development of methodology must be accompanied by theoretical considerations like the ones made in this study. I hope that the proposal of eyebrow movement metaphorically modulating linguistic meaning at different levels is a useful contribution to this agenda.

Works Cited

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2012. "The Essence of Mirativity." Linguistic Typology 16 (3): 435–485. https://doi.org/10.1515/lity-2012-0017.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67(1): 1–48. doi:10.18637/jss.v067.i01.

Bavelas, Janet B., and Nicole Chovil. 1997. "Faces in Dialogue." In The Psychology of Facial Expression, edited by J. Russell and J. Fernandez-Dols, 334–346. Cambridge: Cambridge University Press.

Bavelas, Janet B., and Nicole Chovil. 2000. "Visible Acts of Meaning: An Integrated Message Model for Language in Face-to-Face Dialogue." Journal of Language and Social Psychology 19 (2): 163–194.

Bavelas, Janet B., and Nicole Chovil. 2006. "Hand Gestures and Facial Displays as Part of Language Use in Face-to-Face Dialogue." In The Sage Handbook of Nonverbal Communication, edited by V. Manusov and M. L. Patterson, 97–115. Thousand Oaks: Sage Publications.

Bavelas, Janet B., and Nicole Chovil. 2018. "Some Pragmatic Functions of Conversational Facial Gestures." Gesture 17 (1): 98–127.

Bavelas, Janet B., and Jennifer Gerwing. 2007. "Conversational Hand Gestures and Facial Displays in Face-to-Face Dialogue". In Social Communication, edited by K. Fiedler, 283–308. New York: Psychology Press.

Bavelas, Janet B., Jennifer Gerwing, and Sara Healing. 2014a. "Hand gestures and facial displays in conversational interaction." Oxford Handbook of Language and Social Psychology, edited by T. Holtgraves, 111–130. New York: Oxford University Press.

Bavelas, Janet B., Jennifer Gerwing, and Sara Healing. 2014b. "Including Facial Gestures in Gesture-Speech Ensembles." From Gesture in Conversation to Visible Action as Utterances, edited by M. Seyfeddinipur and M. Gullberg, 15–34. Amsterdam: John Benjamins.

Birner, Betty J., and Gregory Ward. 1998. Information Status and Noncanonical Word Order in English. Amsterdam/Philadelphia: John Benjamins.

Boersma, Paul, and David Weenink. 2018. Praat: Doing Phonetics by Computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from http://www.praat.org/.

Bross, Fabian. 2020. The Clausal Syntax of German Sign Language. A Corpus-Based Approach. Berlin: Language Science Press.
Bross, Fabian, and Daniel Hole. 2017. "Scope-Taking Strategies and the Order of Clausal Categories in German Sign Language." Glossa: A Journal of General Linguistics 2 (1): 1–10. doi: https://doi.org/10.5334/gjgl.106
Carletta, Jean, Amy Isard, Stephen Isard, Jacqueline C. Kowtko, Gwyneth Doherty-Sneddon, and Anne H. Anderson. 1997. "The Reliability of a Dialogue Structure Coding Scheme." Computational Linguistics 23: 13–31.
Cecchetto, Carlo. 2012. "Sentence Types." Sign Language. An International Handbook, edited by R. Pfau, M. Steinbach, and B. Woll, 292–315. Berlin: De Gruyter Mouton.
Chovil, Nicole. 1991a. "Discourse-Oriented Facial Displays in Conversation." Research on Language and Social Interaction 25: 163–194.
Chovil, Nicole. 1991b. "Social Determinants of Facial Displays." Journal of Nonverbal Behaviour 15: 141–154.
Cienki, Alan, and Cornelia Müller (eds.). 2008. Metaphor and Gesture. Amsterdam: John Benjamins.
Darwin, Charles. 1872. The Expression of Emotion in Man and Animals. London: John Murray.
Eibl-Eibesfeldt, Irenäus. 1972. "Ritual and Ritualization from a Biological Perspective." Non-verbal Communication, edited by R.A. Hinde, 297–312. Cambridge: Cambridge University Press.
Ekman, Paul. 2003. Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Times Books/Henry Holt and Co.
Flecha-García, María Luisa 2006. Eyebrow Raising in Dialogue: Discourse Structure, Utterance Function, and Pitch Accents." PhD thesis. University of Edinburgh.
Flecha-García, María Luisa. 2010. "Eyebrow Raises in Dialogue and Their Relation to Discourse Structure, Utterance Function and Pitch Accents in English." Speech Communication 52(6): 542–554.
Gast, Volker, and Johan van der Auwera. 2011. "Scalar Additive Operators in the Languages of Europe." Language 87(1): 2–54.
Guaïtella, Isabelle, Benoît Lagrue, and Chistian Cavé. 2009. "Are Eyebrow Movements Linked to Voice Variations and Turn-Taking in Dialogue"? Lang. Speech 52: 207–222.
Halliday, M.A.K. 1994. An Introduction to Functional Grammar. London: Edward Arnold.
Hamblin, Charles L. 1973 "Questions in Montague English." Foundations of Language 10 (1): 41–53. http://www.jstor.org/stable/25000703.
Herrmann, Annika. 2013. Modal and Focus Particles in Sign Languages. A Cross-Linguistic Study. Berlin: De Gruyter Mouton.

Herrmann, Annika, and Markus Steinbach, eds. 2013. Nonmanuals in Sign Languages. Amsterdam: John Benjamins.
Kim, Jeesun, Erin Cvejic, and Chris Davis. 2014. “Tracking Eyebrows and Head Gestures Associated with Spoken Prosody”. Speech Communication 57: 317–330.
Kimmelmann, Vadim, Alfarabi Imashev, Medet Mukushev, and Anara Sandygulova. 2020. “Eyebrow Position in Grammatical and Emotional Expressions in Kazakh-Russian Sign Language: A Quantitative Study.” PlosOne. DOI: https://doi.org/10.1371/journal.pone.0233731.
Kipp, Michael. 2001. “Anvil – A Generic Annotation Tool for Multimodal Dialogue.” Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech): 1367–1370.
Kratzer, Angelika. 1978. Semantik der Rede. Kontexttheorie, Modalwörter, Konditionalsätze. Königstein: Scriptor.
McNeill, David. 1992. Hand and Mind. What Gestures Reveal about Thought. Chicago: University of Chicago Press.
Patil, Indrajeet 2021. Visualizations with Statistical Details: The ‘ggstatsplot’ Approach. Journal of Open Source Software 6 (61): 3167, doi:10.21105/joss.03167.
Pendzich, Nina-Kristin. 2020. Lexical Nonmanuals in German Sign Language. Berlin: de Gruyter Mouton.
Pfau, Roland, and Josep Quer. 2010. “Nonmanuals: Their Grammatical and Prosodic Roles.” Sign Languages, edited by D. Brentari, 381–402. Cambridge: Cambridge University Press.
Sauter, Disa A. 2017. “The Nonverbal Communication of Positive Emotions: An Emotion Family Approach.” Emotion Review 9 (3): 222–234.
Wells, John C. 2006. English Intonation PB and Audio CD: An Introduction. Cambridge University Press.
Wickham, Hadley (2016). ggplot2: Elegant Graphics for Data Analysis. New York: Springer Verlag.
Wierzbicka, Anna. 2000. “The Semantics of Human Facial Expression.” Pragmatics & Cognition 8 (1): 147–183.
Zeshan, Ulrike. 2004. “Interrogative Constructions in Signed Languages.” Language 80: 7–39.

Absolutely and Totally in Present-Day Spoken British English

1 L N R O D 0 D O H HÏ QR G Y 6 D U N « W D 0 D O £

Charles University, Prague, Czech Republic

[a]nikolamaleckova@seznam.cz, [b]marketa.mala@pedf.cuni.cz

Abstract: The paper explores the various functions of the adverbs absolutely and totally in present-day informal spoken British English. It shows that the adverbs, which were originally semantically close, differ markedly across the three dimensions that the analysis focuses on – sociolinguistic (gender and age distribution), syntactic, and semantic/pragmatic. While the adverbs function most frequently as degree modifiers, they both appear to be following the same trajectory of change from an intensifier, via a stance adverbial to a discourse (response) marker; however, they are currently at different stages of the process of grammaticalization.

Keywords: intensification; stance adverbial; response marker; grammaticalization; Spoken British National Corpus 2014

1. Introduction

In spoken British English, intensification rate has increased during the last two decades: "Male and female speakers across all age groups and all social classes" have been found to use "significantly more intensifiers in 2014 than in 1994" (Fuchs 2017, 355). Underlying the increase in use of individual intensifiers, however, there may be different patterns of use, in sociolinguistic as well as lexico-grammatical terms. The present study focuses on two semantically close intensifiers, whose frequency of use has almost doubled in spoken British English over the last twenty years – absolutely and totally (see Table 1). At the same time, the difference in frequency between the two intensifiers remains statistically significant,[1] with totally being the less frequent one.

7 K H 6 L u H U L H Q W H W L W O H R I D Q D O \ V L V R I W K H 6 L J Q L ¿ F D Q F H Z D V F D O F X O D W H G X V L Q J W K H & \$ / & W R R O & Y

The increase in frequency of absolutely may be accounted for by its having undergone a development along “a trajectory from degree modifier to emphasizer and to discourse marker” (Aijmer 2020, 143), accompanied by an increase in subjectivity. Given the similar rise in frequency of both adverbs, and the overall lower frequency of totally, our main research question is whether totally, originally also a degree modifier, has come to perform the same range of functions as absolutely in present day spoken British English. To compare the current patterns of use of the two adverbs, we will explore their “syntagmatic contexts”, since “constructions (elements in context) and not individual lexical items are the proper domain of grammaticalization” (Himmelmann 2004, 31), focussing on their collocations and syntactic functions.

The distribution of the two adverbs across gender and age groups will also be discussed briefly. Contrary to general sociolinguistic variationist predictions, absolutely is used more often by older speakers than by younger ones, and “it was first used by male speakers and only later spread to female speakers” (Aijmer 2020, 163). We are interested in whether totally follows the same unusual pattern, which could account for its relatively low frequency in present day British English, with the younger speakers preferring “other more fashionable intensifiers” (ibid., see also Núñez Pertejo and Palacios Martínez 2014, 218).

	BNC1994 (demographic spoken component: 5,014,655 words)			Spoken BNC2014 (11,422,617 words)		
	N	per million words	Chi2 6 L J / H	N	per million words	Chi2 6 L J / H
absolutely	671	133.8	213.4423 < .0001	2,875	251.7	995.6747 < .0001
totally	232	46.3		929	81.3	

Table 1. The frequency of absolutely and totally in the demographically sampled spoken component of the BNC1994 and in the Spoken BNC2014.

The paper is organized in the following way. Section 2 gives an overview of previous research into intensifiers, and absolutely and totally in particular. Section 3 describes the corpus used, and outlines the methodology. The following sections discuss and compare the demographic, syntactic and semantic characteristics of absolutely and totally. The results are summarized and commented on in the concluding section.

2 Using similar demographic parameters to sample the population of British English speakers in the United Kingdom, the demographic spoken component of the British National Corpus (1994) is close in its composition to the Spoken BNC (2014).

2. 3URYLRXV 5HVHDUFK

As observed by Bolinger (1972, 18), "[d]egree words afford a picture of fevered invention and competition that would be hard to come by elsewhere, for in their nature they are unsettled." These properties have granted intensifiers substantial attention over an extended period of time, from a variety of perspectives, including diachronic studies (e.g. Méndez-Naya 2008; Tagliamonte and Roberts 2005), synchronic studies from the point of view of age, gender and social class distribution (e.g. Aijmer 2020; Beltrama and Staum Casasanto 2017; Fuchs 2017; Núñez Pertejo and Palacios Martínez 2014; Xiao and Tao 2007), geographical variation (e.g. Schweinberger 2021; Wagner 2017), semantic preference and pragmatic functions (e.g. Beltrama 2018; Partington 2004), or their use in specific settings, such as the dialogue in television series (Aijmer 2016; Quaglio 2009; Tagliamonte and Roberts 2005). Moreover, "[the] continuous waxing and waning of forms in this domain is particularly intriguing as their changing nature predestines intensifier systems to be an ideal case for testing mechanisms of change" (Schweinberger 2021, 233; see also Aijmer 2011, 2018; Bordet 2017; Breban and Davidse 2016; Hessner and Gawlitzek 2017; Macaulay 2006; Partington 1993; Tao 2007). Several studies have been devoted specifically to absolutely and/or totally (e.g. Aijmer 2020, 2016, 2011; Beltrama 2018; Bordet 2017; Carretero 2010; Hessner and Gawlitzek 2017; Núñez Pertejo and Palacios Martínez 2014; Tao 2007; Wagner 2017).

Absolutely and totally are "amplifiers", i.e. intensifiers which "scale upwards from an assumed norm" (Quirk et al. 1985, 445), namely "maximizers", "which can denote the upper extreme of the scale" (ibid., 590). They can modify adjectives, verbs (including participles), prepositional phrases, and nouns (Aijmer 2011, 161; 2020, 150, Biber et al. 1999, 554). Examples (1) and (2) are from the Spoken BNC2014:[3]

(1) (a) and thi- this counsellor she was absolutely brilliant (S28F 1457)
(b) you see again I'd totally forgotten (S28F 2664)

When used as a modifier, absolutely expresses its original, intensifying meaning "to the fullest degree, entirely, wholly, utterly".[4] As far as the collocates of absolutely are concerned, there is not a clear consensus on the semantic prosody of the adverb. Some studies (Partington 2004, 146; Núñez Pertejo and Palacios Martínez 2014, 224) conclude that the adverb does not display a marked preference for either positive or negative collocates. The results of other studies (Tao 2007, 12) show that the positive collocates significantly outnumber the negative ones. The usual collocates of absolutely are of hyperbolic or superlative nature, e.g. enchanting, splendid (Partington 2004, 146).

3 The examples were excerpted from the Spoken BNC2014; the brackets following each example give the text ID code and the utterance number (The British National Corpus 2014: User manual and reference guide 2018, 48)

4 Oxford English Dictionary. Accessed June 28, 2021. http://www.oed.com/.

Totally is defined by the Oxford English Dictionary as: "in a total manner; wholly, completely, entirely, altogether".[5] The previous studies dealing with totally agree on the fact that it usually collocates with negative items expressing "absence" or "lack of", such as, bald, ignored, uneducated, and also with items expressing "change of state" and "transformation", e.g. different, failed (Partington 2004, 147). According to Bordet (2017, 11), the adjectival collocates are often colloquial (e.g. cool, awesome, lame, psyched), and the adverb is generally used in informal contexts.

When modifying verbs, absolutely and totally function as maximal degree adjuncts which are used to intensify the scalar quality of verbs such as love or hate. This use of absolutely is what Aijmer (2016, 83) calls "correct" use. Modifying other verbs, whose meaning does not imply any scale, such as promise or understand, by the adverb absolutely (and by extension totally) is considered "incorrect" by Aijmer (ibid.), as these verbs cannot appear in hyperbolic or exaggerated meanings. By using intensifying adverbs with non-scalar verbs, speakers emphasize the verb and their own YLHZSRLQW LELG í

Both adverbs "seem to be spreading gradually to new syntactic environments by performing new functions which are more discourse-oriented" (Núñez Pertejo and Palacios Martínez 2014, 210). They were observed to function as epistemic content disjuncts (Quirk et al. 1985, 620) or discourse markers, peripheral to the syntactic structure (Aijmer 2020, 157), commenting on the truth value of the utterance and indicating emphasis ("certainly, definitely; without doubt"[6]) or attitude (2). As "clause-oriented" adverbs, absolutely and totally "can be said to convey stronger subjectivity" (Carretero 2010, 213).

(2) S0688[7]: he was older than me wasn't he?
S0687: >>he was older than you absolutely yeah he looked like he was in his sixties (SJM7 1311)

Absolutely and totally were shown to have evolved to function as freestanding "response markers" (Núñez Pertejo 2013, 212), whose "maximizing or totality meaning is carried over to the response and to the hearer to indicate involvement" (Aijmer 2011, 168; see DOTMDRR í

5 Oxford English Dictionary. Accessed June 28, 2021. http://www.oed.com/.

6 Oxford English Dictionary. Accessed June 28, 2021. http://www.oed.com/.

7 In the Spoken BNC2014 transcripts, the codes starting with S refer to the individual speakers.

(3) S0068: well it does it does
S0150: yeah
S0068: I mean
S0075: yeah
S0150: I know it does (.)
S0068: yeah absolutely (S2PS 141)

According to Biber et al. (1999, 551), such adverbs which "stand alone as structurally unconnected elements, and even as complete utterances," can either represent cases of a ellipsis, when the adverb is related to an omitted element of the previous utterance and is used to avoid unnecessary repetition, e.g. The kitten's gone crazy. No, totally I mean it. Totally and utterly <i.e. The kitten has gone totally and utterly crazy.> (ibid.), or as "reaction signals", expressing agreement or an emphatic response to previous utterances. Tao (2007, 19) notes that in these cases absolutely also serves as a turn-taking device. This change in function, together with a semantic shift towards the epistemic meanings of "certainly, definitely, without a doubt", appears to be connected with an ongoing grammaticalization process (ibid., 9). Tao's study of the adverb absolutely shows that the strong tendency to collocate with positive items strengthens the ability of the adverb to carry the positive and affirmative meaning by itself. "Over repeated encounters and as the combinatorial unit of 'absolutely + X' gains positive semantic strength" to the effect that our mental representation of the adverb changes and the need for the X to co-occur with absolutely disappears (ibid., 16). Consequently, absolutely is capable of expressing agreement by itself as a response signal.

Although the two adverbs seem to have evolved in similar ways, there are likely to be differences between them since "similar semantic propensities alone do not guarantee an identical grammaticalization trajectory" (Tao 2007, 15).

Regarding sociolinguistic variation in the use of intensifiers, earlier studies (for a review, see Hessner and Gawlitzek 2017) have shown that absolutely is commonly used by women and by adults in comparison to teenagers, who often use different means of intensification, such as swear or taboo words (Pertejo and Martinez 2014, 218). The distribution of totally, on the other hand, appears to be balanced among genders and age groups.

3. Material and Method

7 K H F D W H Z U L W W H Q F R U S X V / R Y H D O 2017). The corpus consists of transcripts of recordings gathered between the years 2012 and 2016, and therefore provides the most up-to-date data of spoken British English. It contains 11.5 million words and consists of 1,251 casual informal conversations usually

8 The symbol (.) marks a short pause (up to 5 seconds) in the transcripts.

among family and friends, with the total number of 668 speakers (Spoken BNC2014 Manual 2018, 1). The corpus was accessed online, via the CQP web platform.

The use of absolutely and totally was studied from three points of view: a sociolinguistic and a syntactic perspective, and the collocational profile point of view. In the sociolinguistic part of research, the use of the adverbs in the whole Spoken BNC2014 corpus was correlated with the gender and age of the speaker. Since the gender and age sub-corpora are not of equal sizes, the relative frequencies are reported (see Figure 1).

For the study of the syntactic functions of absolutely and totally, 400 random examples of each adverb were extracted from the Spoken BNC2014. The concordance lines were then analysed manually and a syntactic function was ascribed to each adverb. Three primary syntactic functions were distinguished: a modifier, a stance adverbial, and a response marker. The uses of absolutely and totally classified as modifiers were intensifiers of adjectives, adverbs, verbs, noun phrases or prepositional phrases (cf. Aijmer 2020, 150), whose scope is limited to the adjacent phrase (4). Stance adverbials are not syntactically integrated into the clause structure; their scope extends over the whole clause, and they provide "a comment on the content or style of the entire proposition" (Biber et al. 1999, 764) (5). Absolutely and totally can be regarded as response markers when they function as a "response to previous discourse" (Aijmer 2020, 158). In this function, they may be used to express strong agreement, or as backchannels "together with other elements to signal understanding, commitment and involvement" (ibid.) (6).

(4) (a) yeah I know she's she's brilliant absolutely brilliant (S28F 1166)
(b) then they w- swung totally the other way (S28F 2340)

(5) Asterix and Obelix you've never read them but they were my that was all I read when I was your age (.) I should totally get one

(6) S0068: yeah (.) yeah
S0075: but it would get in the way of what I would want to use Facebook for
S0068: absolutely (.) yeah (.) yeah (S2PS 229)

The collocational patterns of the adverbs are based on the whole Spoken BNC2014. The list of collocates was calculated using the log-likelihood statistics (with the cut-off point set at 6.63), the span was restricted to one position to the right of the adverb. The collocations were manually checked and wrong hits were removed. The semantic prosody of adjectival and verbal collocates was determined by examining the individual

) R U D [W D E V R O H W H R J O L N H O L K R R G U K L J K H U L J Q L J D W D Q H Z Y R H O
$p < 0.01$ (cf. Brezina 2018, 114).

lexemes in isolation. In indeterminate cases, a broader context of the respective concordances was considered.

4. The Frequency and Sociolinguistic Variables

In the Spoken BNC2014, absolutely was found to be almost three times as frequent as totally, with 2,875 and 929 occurrences, respectively (see Table 1 above).

Each of the adverbs shows a different distribution with respect to gender and age (see Figure 1). While the distribution of absolutely varies considerably with regard to gender and age, the use of totally is more evenly distributed. Absolutely is used significantly more often by women than by men ($p < .0001$).[10] In contrast, there is no significant difference in the use of totally between women and men. Among women, the use of absolutely rises significantly with age in the age groups of 35–44, 45–59 and 60+. In the case of men, the only significant rise in use of absolutely appears between the age groups 0–24 and 25–34. The age distribution of totally is different; women's use of totally shows a significant rise only between the first two age groups. Men, on the other hand, show a significant rise in use in the age group of 60+.

The adverbs absolutely and totally are least commonly used by speakers aged 0–24. These results correspond with those of Núñez Pertejo and Palacios Martínez (2014, 218), who point out the differences in the use of -ly adverb intensifiers between adults and teenagers:

> -ly adverbs tend to be more common in formal language and hence do not sit comfortably with the spontaneity of youth interactions; […] teenagers resort more often than adults to taboo and swear words as intensifiers, on many occasions these functioning as markers of in-group solidarity and identity, and possibly also as a sign of rebellion against the adult world and the establishment.

By far the most active users of absolutely are women over the age of 60, which corresponds with the findings of Aijmer (2020, 148).

10 The Bonferroni correction was employed in order to prevent the occurrence of false S R V L W L Y H V L J Q L ¿ F D Q F H W K U H V K R O G Z D V W K X V V H W D W

Figure 1. Distribution of *absolutely* and *totally* with regard to gender and age (ipm = items per million words).

5. Syntactic Functions and Colligations

Both *absolutely* and *totally* function most frequently as modifiers, as shown in Figure 2. *Absolutely* was found to perform the function of a modifier in more than half of all its instances (56%). Most prominently, the adverb colligated with adjectives (7 a), including participial adjectives (7 b), and took on the well-established role of an intensifier, indicating "an endpoint on a scale" (Biber et al. 1999, 555).

(7) (a) and understanding the realities of it all she was *absolutely* fantastic (SRD5 405)
(b) I'd been at work all week and I was *absolutely* knackered (SK8T 358)

In several cases *absolutely* appeared as a modifier of verbs (8). The modified verbs always imply a scale and thus allow for intensification.

(8) I'd hate it I'd *absolutely* hate it but I think it's probably good for me (S5H 251)

Figure 2. Distribution of absolutely and totally across syntactic functions (raw frequencies).

Absolutely frequently modified noun phrases. It preceded either the determiner no (9 a), the pronouns nothing, nobody, anything, or all (9 b, c), or a noun expressing amount, such as plenty, loads, or tonnes (9 d). Generally, the directly modified item referred to quantity.

(9) (a) well I have absolutely no intention of erm switching (S8BQ 278)
(b) and there was nothing left absolutely nothing to show for it (SRVR 106)
(c) well your first year's like you spend a whole year doing absolutely fuck all (SY2Z 351)
(d) which should have been absolutely plenty R I W L P H 6 / (

Modification of adverbs (10 a) and prepositional phrases (10 b) by absolutely proved rare.

(10) (a) she said doctor that worked absolutely wonderfully 6 + / 7
(b) but the Church is absolutely against it (SQWC 1067)

The adverb totally appeared as a modifier in 259 cases out of 400 (64.75%). Its most frequent function was that of a modifier of adjectives, as was the case of absolutely was also found to colligate with participial adjectives (11).

(11) (a) it's like totally ridiculous (SES6 975)
(b) I remember just getting totally slaughtered 6 / 6

Totally was more frequent than absolutely as a modifier of verbs (12 a), adverbs (12 b), and prepositional phrases (12 c). All the modified verbs imply change, and can therefore be intensified. The numbers of adverbs and prepositional phrases are too low to allow any generalization.

(12) (a) they've totally changed the form (SABT 27)
(b) well he totally falsely accused me of being absolutist (SP2X 1287)
(c) I'm totally into that (S23A 2428)

In comparison to absolutely, totally very rarely modified noun phrases; in the present data set, it colligated with the determiner all, the pronoun some, and the noun phrase the other way.

In the role of a stance adverbial, the adverb totally was markedly more frequent than absolutely (90 and 22 instances, respectively) to Finally, a stance adverbial was its second most frequent syntactic function, appearing in 22.5% of instances, while stance adverbials made up only 5.5% of all uses of absolutely, being the least common syntactic role of the adverb. When used as stance adverbials, both absolutely and totally express high levels of conviction or emphasis. In this function the adverbs may be replaced by definitely or certainly, or may be paraphrased as "it is definitely the case that". In the present data, both adverbs appeared as stance adverbials in initial (13 a), medial (13b, as well as final position (13d).

(13) (a) S0015: and I ended up with some vouchers left at the end
S0566: totally totally a waste of money (S94Z 523)
(b) it's totally recording right now (SXEV 877)
(c) well it's my guarantee and absolutely can't lose that because
(d) and it is a scam like totally (SHKF 512)

The role of a response marker was the second most prominent function of absolutely (34.25% of its instances). Totally only appeared as a response marker in 30 cases (7.5%). Both adverbs were used more frequently as backchannels (14) than answers to direct questions (15). Absolutely and totally were used predominantly to express agreement or an affirmative answer, often in combination with other response tokens, such as yes, yeah,

11 The stance adverbial uses of totally and absolutely may be hard to distinguish from the [illegible] doubt that" signalling epistemic meaning could be applied (see Tao 2007, 15), e.g. totally can (.) I get that [illegible] adverbials from response markers.

or oh. In a few cases, however, the adverbs were found to be followed by the adverb and thus serve as emphatic rejections (16). Absolutely appeared in such use in 4 instances, totally appeared with not only once.

(14) S0515: oh I think it's very important to be very helpful
S0511: absolutely yeah (SKX6 665)

(15) S0561: has yours and –ANONnameF's cycles synced?
S0544: oh yeah totally (SVHN 1405)

(16) S0266: are you gonna climb Mount Everest?
S0270: on a no absolutely not I've looked at it (S7KK 34)

The frequent use of absolutely as a response marker supports the previous findings about its undergoing a process of grammaticalization (Tao 2007). The original meaning of wholeness or completeness is disappearing, and the adverb comes to mean "yes", "certainly", "sure". The underlying sense of maximization, originally associated with intensification of adjectives and verbs. is now being used to express agreement or strong speaker involvement. Totally can also be used in this manner; it comes to bear the same affirmative meanings as absolutely. This signals that it is undergoing grammaticalization as well. Yet the use of totally as a response marker is rather infrequent, while the adverb performs the function of stance adverbial more frequently, which suggests that in the case of totally the process of grammaticalization is less advanced.

The category of "other" in Figure 2 refers to those instances in our data set that could not be classified due to incompleteness of the utterances.

6. Collocational Patterns

The most common collocates of the adverb absolutely were adjectives, verbs, noun phrases, and interjections. Table 2 shows the distribution of the types and tokens of the adjectives and verbs modified by absolutely with regard to their semantic prosody.

Semantic Prosody	Adjectival collocates		Verbal collocates	
	Types	Tokens	Types	Tokens
Favourable	33	460	4	86
Unfavourable	62	328	13	59
Neutral	24	131	9	52
TOTAL	119	919	26	197

Table 2. Collocates of absolutely (based on the whole Spoken BNC2014).

Among the significant adjectival collocates of the adverb absolutely, there were more favourable tokens but at the same time more unfavourable types of adjectives. Thus, there is less variation among the positive adjectives but their association with absolutely is arguably more well-established. On the other hand, the adjectives with negative semantic prosody are diverse, but clearly the collocations are not as widespread, and their individual frequencies are overall lower.

The most frequent adjectival collocates of absolutely are the words fine, brilliant, and amazing, which appear with the absolute frequency of 97, 58 and 49, respectively. The typical adjectival collocate of absolutely bears the meaning of extreme or superlative quality (17 a), for example, gorgeous, fantastic, stunning. The majority of the adjectival collocates (64 out of the 119 types) in our data expressed subjective evaluation and carried a high emotional load (17 b), e.g. ridiculous, delicious, amazing, awful.

(17) (a) cos we were looking at flowers in an absolutely enormous S R W 6 / 0 %
(b) Hitchhiker's Guide to Galaxy which was absolutely terrible 6 3 0 /

Another prominent semantic group of adjectives intensified by absolutely comprises adjectives describing some internal or external characteristic or emotional state: violent, adamant, horrified, appalled. There also appears a homogenous semantic group of adjectival collocates of absolutely referring to a physical state – all are participial adjectives which refer either to someone being drunk (smashed, fucked, slaughtered, battered) (18 a) or tired (exhausted, knackered) (18 b).

(18) (a) his partner she was like absolutely blottoed (S6A5 335)
(b) I just had no energy I was just like absolutely shattered (SMZV 304)

Absolutely functions as a modifier of verbs less frequently than as a modifier of adjectives, and thus the number of its significant verbal collocates is lower – 26 types. The two verbs most prominently collocating with absolutely are love and hate, which appear with the absolute frequency of 75 and 26, respectively. The synonyms of these verbs, adore and loathe, also appear among the ten most common verbal collocates. As in the case of adjectives, the verbs modified by absolutely are expressive and often colloquial, e.g. chuck, piss down, reek, ruin, howl. The verbs may describe both physical actions (19 a) and cognitive processes (19 b).

(19) (a) when we got there it was absolutely pouring from rain (SWRQ 168)
(b) but yeah I absolutely agree with you (SVXP 424)

Table 2 demonstrates that verbal collocates of absolutely display the same pattern of semantic prosody as the adjectival collocates: there are fewer types of favourable verbal collocates but they outnumber the negative verbal collocates in terms of tokens. Notably, of 75

the total 86 favourable tokens are instances of the verb love. The rest of the favourable collocates are therefore much less frequent. As in the case of the adjectives, there exists a greater variety of the unfavourable or neutral verbs modified by the adverb. The collocation absolutely + verb thus seems to be prone to be unfavourable or neutral rather than favourable – in which case there is a strong preference for the collocation absolutely love.

As shown in Section 5, when modifying noun phrases, absolutely may be followed by a pronoun (nothing, everything, anything, none, nobody, all), a determiner (no) or a noun. There are in total 5 significant nominal collocates, all expressing some quantity: ages, loads, tonnes, hours, years, which further supports clear semantic preference of absolutely to collocate with noun phrases that express quantity or extent.

As noted in the previous section, absolutely in the role of a response marker is commonly accompanied by other non-clausal units, such as interjections (oh), fillers (er, erm), and response forms (yeah). In the majority of cases, the adverb is used to express agreement, and is therefore combined with yeah, yes, or functionally congruent interjections, such as aye, mhm, uhu. Even when absolutely collocates with no, the overall meaning of the utterance is actually agreement with the previous utterance (20).

(20) S0262: just goes to show how political education is isn't it?
S0301: yeah absolutely no absolutely that's absolutely right (SA6K 214)

The overall semantic prosody of the adverb absolutely cannot be determined unequivocally due to the uneven distribution of favourable and unfavourable types and tokens. There is a common pattern of the adjectival and verbal collocates of absolutely, where the favourable items are more frequent as tokens but fewer as types. This suggests that while there is less variation of the favourable items, they are more strongly established in the speakers' lexicons and appear more frequently in speech. On the other hand, it seems that absolutely also lends itself easily to be used with a array of unfavourably or neutrally evaluative items.

The word classes most frequently collocating with the adverb totally are adjectives, verbs, prepositional phrases, and interjections. Table 3 summarizes the distribution of the adjectival and verbal collocates with regard to semantic prosody.

Semantic Prosody	Adjectival collocates		Verbal collocates	
	Types	Tokens	Types	Tokens
Favourable	8	27	2	24
Unfavourable	25	81	9	60
Neutral	17	135	13	93
TOTAL	50	243	24	177

Table 3. Collocates of totally (based on the whole Spoken BNC2014).

In the case of totally, most of the adjectival collocates were either unfavourable or neutral in nature – favourable items were the least frequent both by type and token. By far the most frequent collocate is the adjective different, which appears with the absolute frequency of 93. The adjectives modified by totally were overall less expressive and hyperbolic than in the case of absolutely (21). Most of the adjectives referred either to an internal or an external characteristic of a thing or a person (refined, committed, opposite, honest) or expressed evaluation (wrong, normal, insane, awesome). There can also be identified a semantic group of adjectives expressing "a lack of something": empty, bald, nonsensical, devoid, insane.

(21) It means that what they're doing is totally wrong (S7KK 698)

The neutral or negative semantic prosody of totally is retained when the adverb modifies verbs. There are only two verbs collocating with totally – agree and understand – which can be classified as favourable. In general, the collocates of totally are common neutral verbs, which can hardly be considered expressive or exaggerated (22). The only semantic group identified in the list of collocates are verbs relating to cognitive processes: forget, agree, know, ignore, focus, imagine. With the absolute frequency of 29, forget is the most frequent verbal collocate of totally.

(22) I'd totally forgotten that we're still recording (SB4D 1871)

Totally is quite frequently followed by a prepositional phrase, most prominently by out of, e.g. totally out of the blue, totally out of proportion, totally out of control, totally out of line.

When functioning as a response marker, totally behaves in the same manner as absolutely, as it usually collocates with positive interjections and with fillers, such as aye, yep, yes, mm, and oh. By far the most frequent response form collocating with totally is yeah.

The semantic prosody associated with totally is much clearer than that of absolutely as it displays a general preference for collocates with unfavourable or neutral meanings, unless it appears as a response signal, in which case it bears the positive semantic prosody of an agreement.

7. Discussion and Conclusions

Quoting Tao's (2007, 23) observation that "degree adverbs are among the most active classes of words in terms of undergoing change", Aijmer (2020, 161) asks "whether grammaticalization can explain the semantic and pragmatic developments of absolutely".

This is in contrast with the findings of Bordet

Our results support her conclusions about the expansion of absolutely related to the three VWDJHVDPPDWLFEHKDLPSWORQQí QDPHVWKLQFUHDVH frequency over time, new syntactic functions (34.3% of instances of absolutely function as response markers, and 5.5% as stance adverbials), and "semantic-pragmatic expansion", i.e. the expression of stance and emphasis.

The comparison of syntactic functions performed by absolutely and totally has shown that both adverbs appear most frequently as modifiers of adjectives, which is their original degree-indicating role, but their distribution across other syntactic functions differs. Besides absolutely, the adverb totally can express agreement and be used as a response marker, but this function is rather rare (7.5% of occurrences of totally). Totally is used more frequently as an epistemic marker, indicating the speaker's certainty ("definitely" or "without a doubt"). The difference in the extent of expansion into new syntactic domains suggests that while the two adverbs appear to be following the same trajectory of change from a degree modifier, via a stance adverbial to a discourse (response) marker, they are at different stages of the process of grammaticalization. In comparison with absolutely, totally has also increased in frequency, but has been more restricted in its spread into new syntactic functions, and neither has it fully undergone the semantic-pragmatic expansion that can be observed in the case of absolutely. Its ability to serve as a stand-alone discourse marker emphasizing full agreement is limited.

The differences between the two adverbs may be related to some factors suggested by the analysis of the sociolinguistic variables of age and gender, and of the collocational patterns. Each of the adverbs displays a different pattern of distribution with respect to gender and age. While absolutely is used more commonly by women, in the case of totally there is virtually no gender difference, and the use of the adverb is more balanced throughout the age groups as well. Absolutely, on the other hand, has a tendency to increase in frequency with the age of the speaker, particularly among female speakers. Aijmer (2020, 163) concludes that "older (female) speakers may use absolutely consciously and with higher frequency because of its expressive and hyperbolic qualities".[13] The role of expressiveness as a factor contributing to the spread of the adverb appears to be borne out by the collocational analysis.

The analysis of collocations has shown that the two adverbs differ in their semantic prosody and preference. There has not been a consensus on the semantic prosody of absolutely. The present findings support those of Partington (2004) and Núñez Pertejo and Palacios Martínez (2014), namely that absolutely does not show a preference for either favourable or unfavourable items. It has been shown that while the favourable collocations

13 On the association between linguistic expression of emotion and gender, see e.g. Coates (2015) or Fuchs (2017); Tagliamonte and Roberts (2005: 289) suggest that "[the] correlation RIZRPHQZLWKQWHQVLYHDGYHUEVPD\UHIOHFWZRPHQ¶VPRUHH[SUHVVLYHUROHLQ language than men".

seem to be more well-established and more frequently used by speakers, there is a greater variety of the unfavourable items that may collocate with the adverb. The collocations of absolutely were also found to be of hyperbolic and superlative nature, e.g. absolutely brilliant, absolutely hate. The semantic prosody of totally is easier to determine as the adverb frequently collocates with either unfavourable or neutral items. It seems likely that the high frequency of absolutely (compared to totally) may be related to the general "tendency for use of intensifiers with emotional language" (Tagliamonte and Roberts 2005: 290). In contrast, across various word classes, the collocates of totally are usually neutral, inexpressive terms. There is little overlap between the collocates of both adverbs in the functions of a modifier and stance adverbial. For instance, there are only 13 adjectives that can collocate with both absolutely and totally, e.g. honest, ridiculous, sure, fine, normal.

In this paper, we hope to have provided a comprehensive comparison of the present-day uses of the adverbs absolutely and totally in spoken British English. It has been shown that the adverbs, which were originally semantically close (meaning "entirely" or "wholly"), differ markedly across the three dimensions that were the focus of the present analysis – sociolinguistic, syntactic, and semantic. The semantic prosody of absolutely is more equivocal than that of totally, whose collocates are generally negative or neutral. Absolutely, on the other hand, displays a semantic preference for expressive, hyperbolic and colloquial collocates, while the collocates of totally tend to be stylistically unmarked and inexpressive. These differences may have impact on the overall higher frequency of absolutely and its association with particular groups of speakers (older female speakers), who may be using it consciously due to its expressiveness. This contrasts with the neutral collocates and more balanced gender and age distribution of totally. Both adverbs appear to be following the same grammaticalization path, but while absolutely has expanded not only in terms of frequency, but also syntactically, semantically and pragmatically (as attested by its widespread use as a response discourse marker), totally seems to be at a lower stage of the process of grammaticalization, being used predominantly as a modifier and a stance adverbial.

With the intensification rate recently increasing "across the board" in British English (Fuchs 2017, 346), it would be interesting to explore the potential of individual intensifiers, such as completely or definitely, to acquire new syntactic functions and expand into new semantic-pragmatic domains, perhaps following a grammaticalization trajectory similar to that of absolutely and totally.

Works Cited

Aijmer, Karin. 2011. "Are You Totally Spy? A New Intensifier in Present-Day American English." In Marqueurs Discursifs Et Subjectivité, edited by Sylvie Hancil, 155–172. Rouen: Universités De Rouen and Havre.

Aijmer, Karin. 2016. "'You're Absolutely Welcome, Thanks for the Ear': The Use of Absolutely in American Soap Operas." Nordic Journal of English Studies 15 (2): 78–94.

Aijmer, Karin. 2018. “That’s Well Bad. Some New Intensifiers in Spoken British English.” In Corpus Approaches to Contemporary British English, edited by Vaclav Brezina, 5 R E E L H / R Y H $ D M Q G H U D U L Q < R U 1 H Z Q G / R Q G R Q 5 R X W O
Aijmer, Karin. 2020. “That’s Absolutely Fine. An Investigation of Absolutely in the Spoken BNC2014.” In Corpora and the Changing Society: Studies in the Evolution of English, edited by Paula Rautionaho, Arja Nurmi, and Juhani Klemola, 143–167. Amsterdam and Philadelphia: John Benjamins.
Beltrama, Andrea. 2018. “Totally Between Subjectivity and Discourse. Exploring the Pragmatic Side of Intensification.” Journal of Semantics 35: 219–261.
% H O W $ U D G U D H Q G 6 X 6 W D & P V D V D Q W R W D O O \ Younger: Intensification at the Socio-Semantics Interface.” Journal of Sociolinguistics 21 (2): 154–182.
% L E ' H U J 6 W 4 K D Q ¨ V H R Q U H F K V & Q Q D G Z D U L H J D Q 1999. Longman Grammar of Spoken and Written English. D U O R Z / R Q J P D Q
Bolinger, Dwight. 1972. Degree Words. The Hague: Mouton.
% R U G ; H F O H ³) U R P J X U W R [L F D O , Q W H Q V L ¿ L F D W L R Q V R X U G Y J The Renewal and Recycling of Intensifiers in English. A Case-Study of Very, Really, So and Totally.” Lexis (Online) 10. Accessed July 19, 2019. http://journals.openedition.org/lexis/1125.
Brezina, Vaclav. 2018. Statistics in Corpus Linguistics: A Practical Guide. Cambridge: Cambridge University Press.
Carretero, Marta. 2010. “You’re Absolutely Right! A Corpus-Based Contrastive Analysis of Absolutely in British English and Absolutamente in Peninsular Spanish, With Special Emphasis On the Relationship Between Degree and Certainty.” Languages in Contrast 10 (2): 194–222.
Coates, Jennifer. 2015. Women, Men and Language: A Sociolinguistic Account of Gender Differences in Language / R Q G R Q 5 R X W O H G J H
' X ã N R Y L E X ã 6 H O 0 O X Y Q L F H V R X þ D V Q p D Q J O L þ W L Q \ Q D Edition. Accessed January 28, 2020. Praha: Ústav anglického jazyka a didaktiky, FF UK. http://www.mluvniceanglictiny.cz
Fuchs, Robert. 2017. “Do Women (Still) Use More Intensifiers Than Men? Recent Change in the Sociolinguistics of Intensifiers in British English.” International Journal of Corpus Linguistics 22 (3): 345–374.
Hessner, Tanja, and Ira Gawlitzek. 2017. “‘Totally or Slightly Different?’ A Spoken BNC2014-Based Investigation of Female and Male Usage of Intensifiers.” International Journal of Corpus Linguistics 22 (3): 403–428.
+ L P P H O P D Q Q 1 L N R O D X V ³ / H [L F D O D Q G G U D P D W L F D S S U R D Y W W R Q or Orthogonal?” In What Makes Grammaticalization? A Look Form Its Fringes and Its Components, edited by Walter Bisang, Nikolaus P. Himmelmann, and Björn Wiemer, 21–42. Berlin and New York: Mouton De Gruyter.

/ R Y H U E D & L H ' H U P H E U \ Q G U H D Z U G D H F % D U H] D Q G R Q \ F H Q H U \ “The Spoken BNC2014: Designing and Building a Spoken Corpus of Everyday Conversations.” International Journal of Corpus Linguistics 22 (3): 319–344.

Macaulay, Ronald. 2006. “Pure Grammaticalization: The Development of a Teenage Intensifier.” Language Variation and Change 18 (3): 267–283.

Núñez Pertejo, Paloma. 2013. “From Degree Adverb to Response Token: ‘Absolutely’ L Q / D W H G H U Q G R Q W H P % R U D Q S K U L (Q J O L V K ” Anglia – Zeitschrift für englische Philologie 131 (2): 207–235.

Núñez Pertejo, Paloma and Ignacio M. Palacios Martínez. 2014. “That’s Absolutely Crap, Totally Rubbish: The Use of the Intensifiers Absolutely and Totally in the 6 S R N H Q J X H % U L W L V K ” Functions of Language 21 (2): 210–237.

Oxford English Dictionary. 2019. Oxford: Oxford University Press. Accessed June 28, 2021. Available online at http://www.oed.com/.

3 D U W L Q J W R Q & R U S X V L G H Q W - H Q V L fiers.” In Text and Technology: In Honour of John Sinclair, edited by Mona Baker, Gill Francis, and Elena Tognini-Bonelli, 177–192. Amsterdam and Philadelphia: John Benjamins.

Partington, Alan. 2004. “Utterly Content in Each Other’s Company. Semantic Prosody and Semantic Preference.” International Journal of Corpus Linguistics 9 (1): 131–156.

Quaglio, Paulo. 2009. Television Dialogue. The Sitcom Friends vs. Natural Conversation. Amsterdam and Philadelphia: John Benjamins.

4 X L U N G R O I H U H Q H D U W Y L N Comprehensive Grammar of the English Language. / R Q G R Q / R Q J P D Q

Schweinberger, Martin. 2021. “Using Intensifier-Adjective Collocations to Investigate Mechanisms of Change.” In Variation in Time and Space, edited by Anna ý H U P i N R Y i D Q G 0 D U N p W D 0 D O i ± % H U O L Q ' H

Tagliamonte, Sali, and Chris Roberts. 2005. “So Weird; So Cool; So Innovative: The Use of Intensifiers in the Television Series Friends.” American Speech 80 (3): 280–300.

Tao, Hongyin. 2007. “A Corpus-Based Investigation of Absolutely and Related Phenomena in Spoken American English.” Journal of English Linguistics 35 (1): 5–29.

The British National Corpus 2014: User Manual and Reference Guide. Version 1.1. 2018. Accessed June 8, 2021. Available online at http://corpora.lancs.ac.uk/bnc2014/doc/bnc2014manual.pdf.

Wagner, Susanne. 2017. “Totally New and Pretty Awesome: Amplifier-Adjective Bigrams in Glowbe.” Lingua 200: 63–83.

Xiao, Richard, and Hongyin Tao. 2007. “A Corpus-Based Sociolinguistic Study of Amplifiers in British English.” Sociolinguistic Studies 1 (2): 241–273.

Sources and Tools

Spoken BNC2014: The Spoken British National Corpus 2014. Accessed June 8, 2021. Available online at https://cqpweb.lancs.ac.uk/.

& Y U þ9H-N O D Y Calc: Corpus Calculator Praha: FF UK. Accessed June 8, 2021. Available online at https://www.korpus.cz/calc/.

CQP Web. Accessed June 8, 2021. Available online at https://cqpweb.lancs.ac.uk/.

Perception and Production of Geminate Timing in Hungarian Voiceless Stops

Tilda Neuberger

Hungarian Research Centre for Linguistics, Budapest, Hungary

neuberger.tilda@nytud.hu

Abstract: Hungarian is a language that expresses semantic differences using contrastive consonant length. Previous research in many languages confirmed that duration is the most important acoustic correlate of the singleton-geminate contrast. The present study aims to examine the acoustic and perceptual correlates which distinguish single stops from their geminate counterparts in Hungarian. The main question is how closure duration contributes to the perception and production of the length contrast of voiceless stops? Results confirmed that closure duration is the main acoustic attribute that distinguishes between singletons and geminate stops in Hungarian, and it is a sufficient cue for discriminating them in perception as well. Analysis of the relationship between consonant and preceding vowel duration did not support the strategy of temporal compensation. Findings are supposed to explore the relationship between the acoustic and perceptual domains and shed light on the primary/secondary acoustic features of consonant length opposition in Hungarian.

Keywords: geminate; stop consonant; duration; perception; Hungarian

1. Introduction

Gemination has been widely investigated in many languages. Production studies confirmed that the primary acoustic cue for the singleton-geminate (S-G) contrast is the duration (e.g., Ham 2001; Ridouane 2010). Moreover, it was suggested that in the case of complex speech sounds, such as stops or affricates, a detailed examination is needed to find out which portions of the internal structure (closure duration [CD], voice onset time [VOT], etc.) are targeted by phonological lengthening (Pycha 2009). In the

of stops, CD proved to be the most important acoustic attribute of the S-G contrast; however, in several languages the duration of the previous vowel (V1) also seemed to contribute to the opposition (Ridouane 2010).

In many languages, V1 is realized longer before singletons (e.g. Italian: Esposito and Di Benedetto 1999; Swedish: Hassan 2003). This can be explained by temporal compensation at the segmental level. This means that each VC or CV sequence is time-balanced: a shorter duration of one segment leads to a longer duration of an adjacent segment, and the increased duration of a segment is compensated by a shortening of another segment (Hassan 2003; Khattab 2007; Issa 2015). However, such an inverse relationship between V1 and G duration is not clearly supported in all languages (e.g., Iraqi Arabic; Hassan 2003; Lebanese Arabic: Khattab and Al-Tamimi 2008), and there is even research that has found the opposite trend, increasing V1 duration in the geminate environment (e.g., Persian: Hansen 2004; Japanese: Idemaru and Guion 2008; Fujimoto et al. 2015; Finnish: O'Dell and Malisz 2016; Hungarian: Deme et al. 2019). In addition, some studies did not find any difference in these terms (e.g., Polish: Rojczyk and Porzuczek 2019). The various kinds of relationship between V1 and G duration might be traced back to language typological reasons (e.g., differences between mora-timed and syllable-timed languages; see Ham 2001; Idemaru and Guion 2008).

Perception studies investigated the role of durational parameters, such as CD, the ratio of consonant duration to vowel/mora/word duration in length categorization (e.g., Hankamer et al. 1989; Amano and Hirata 2010; Idemaru and Guion-Anderson 2010). They found that not only absolute but also relative durations proved to be adequate perceptual cues in the distinction of single and geminate stops, especially when speech/articulation rate is taken into account.

2. Gemination in Hungarian

Length is a phonologically relevant feature in the consonant system of Hungarian, i.e., geminate and singleton consonants are contrastive, e.g., ép > H • S @ µ K H ép p > H • S • @ µ U L J K W Q R Z ¶ $ O O F R Q V R Q D Q W V L Q + X Q J is also phonemic, and both short and long vowels can occur before or after geminate consonants, e.g. hallani > K % O • % Q L @ á s µ V W K % K O • D U ¶ á @ a m µ K H D O • L Q J @ µ V W D W H ¶ á l ¶ á s \ > H K D O • U @ µ M R E ¶

The distribution of geminates is restricted in Hungarian: geminates cannot stand word-initially or next to another consonant. "In other words, the occurrence of geminates is only possible (i) intervocalically (e.g. állat µ D Q L á t Q ¶ F K L Q á l D t ¶ µ $ W W L O D stands') and (ii) utterance finally (i.e., before a pause) if preceded by a vowel (e.g. áll µ V W D Q G ¶ µ F K L Q ¶ ´ 6 L S W i U D Q G * U i F] L

According to the traditional view, it is assumed that geminates must surface as short if they are flanked by another consonant on either side, (this process is called G H J H P L Q D W L R Q 6 L S W i U D Q G * U i F] L ' H P H H W D O

degeminated stop duration was more than 88% of singletons, i.e., they were even shorter than singletons, and degeminated stops patterned with singletons in two-term clusters (Deme et al. 2019).

Three types of geminates can be distinguished based on their abstract phonological representation:

- Underlying/lexical geminate: part of the phonemic inventory of the language, for instance: sok > U R N @ µ P D Q \ ¶ sokk > U R N • @ µ V K R F N ¶
- Derived true geminate: results from some assimilation processes (e.g., voicing assimilation, v-assimilation), for instance: kalap + -val (INSTR) > kalappal > N % O % S • % O @ µ Z L W K K D W ¶
- Fake geminate: merged sequences of identical consonants arising through morpheme concatenation, for instance: zseb + -ben (INE) > zsebben > d - E • - Q @ µ L Q S R F

The comparison between the three types of geminates revealed that the acoustic correlates of fake geminates differed from underlying or derived true geminates to some extent (Neuberger 2015).

Geminates in Hungarian have a relatively low functional load. Consonant length as a phonological feature most often distinguishes minimal pairs whose one or both members are loanwords (e.g., kasza > N % V % kasszа µ V > F N \ W K • P ¶ @ µ F D V K ¶ names, or contain derived geminates (not underlying ones, e.g., vasal > Y % U % O @ ironing' : vassal > Y % U • % O @ µ Z L W K L U R Q ¶ vas + -val (INSTR)).

3. Aims of the Present Study

The aim of the present study is to examine the acoustic and perceptual correlates which distinguish single /p, t, k/ stops from their geminate counterparts in Hungarian. We hypothesized that closure duration plays the most important role in the distinction between singletons and geminate stops in Hungarian, and is sufficient cue for discriminating them in perception as well. It is also assumed that secondary cues for the distinction might be manifested in vocalic context duration. A comparison of production data and perception test results is made to explore the relationship between the acoustic and perceptual domains.

4. Acoustic Study

Firstly, in order to examine the role of temporal cues in the distinction between singletons and geminates in Hungarian, we conducted an acoustic analysis of intervocalic stops produced in spontaneous speech. The main questions here were: how do Hungarian speakers distinguish singleton stops from geminates acoustically in everyday speech? More precisely, how much durational difference can be found between voiceless singleton and geminate stops (based on closure duration, voice onset time)? How does vowel

V ¶

duration contribute to the consonant length distinction? We investigated several durational parameters related to the target consonant and surrounding vowels.

4.1 Method

Ten adult males (ages 21–29, mean = 24.7 years) with normal voice quality and no reported history of speaking or hearing disability participated in this study. All participants are monolingual, native speakers of standard Hungarian. Data were drawn from the BEA database (Neuberger et al. 2014). While previous studies have generally examined the phonetic realization of consonant length in controlled material (e.g., Pycha 2009; Demeet al. 2019), the present research investigates spontaneous speech. In this speech type, the durational overlap is expected to be more significant, with a lower singleton-geminate ratio than in read word or sentence list (see Khattab 2007). Recordings were obtained in the same sound-proof room, with AT4040 microphones, using GoldWave sound editing software (sampling at 44.1 kHz, storage: 16 bits, 86 KBytes/s, mono). Participants talked about their job and free-time activities.

The data set contained 590 manually segmented stop consonants (260 Ss and 330 Gs) in intervocalic positions. Underlying and derived true geminates were selected, but fake geminates were excluded from the analysis. The following parameters were measured in Praat 6.1 (Boersma and Weenink 2020): CD, VOT, total C-duration (CD+VOT). In the next step, for examining relative consonant duration related to adjacent vowel duration, RQO\ WKH PRVW FRPPRQ YRZHOV LQ WKH PDWHULDO LQ WKH SUHVHQW UHVHDUFK PDWHULDO > % @ > R balanced according to the position (preceding or following). The measurement was extended to the duration of 244 V1 and 237 V2. We measured C/V1 and C/V2 ratios, i.e., the ratios of total consonant duration to the preceding and following vowel duration, as well as G/S ratios, i.e., mean geminate to singleton ratios per speaker.

Statistical analysis was made in R (R Development Core Team 2019) using lmerTest package (Kuznetsova et al. 2017). The mean and standard deviation of the raw data are indicated in the text. For the purposes of the statistical analysis, all absolute durational measurements were log-transformed. Using linear mixed models, we examined the HIIHFW RI µTXDQWLW\¶ DQG µSODFH RI DUWLFXODWLRQ YDULDEOHV DQG µVSHDNHUV¶ ZHUH GHILQHG DV UDQGRP µYRZHO TXDOLW\¶ LQ RXU FDVH > ' @ RU > R-values ZDV H[DPLQHG and corresponding p-values were computed using the Satterthwaite method.

4.2 Results

First, we analysed the data pooled over all consonants. Singletons were realized with an average duration of 96.62 (SD: 19.57) ms, while geminates were realized with an average duration of 140.66 (SD: 28.38) ms. G/S ratio was 1.4, on average (ranging between 1.2–1.6 across speakers). Statistical analysis confirmed that single and geminate stops differed

significantly in terms of total consonant duration (1, F584) = 419.07, p 0 R U H R Y H U place of articulation also had a main effect on consonant duration (2, F584) = 3.903, p = 0.021, but the interaction between the two factors was not significant.

CD also differed significantly between singletons and geminates (Fig. 1): F(1, 548) = 470.10, p 7 K H R Y H U O D S E H W Z H H Q W K H W typically in the duration range between 50 and 120 ms. POA had a main effect on CD: F(2, 548) = 13.89, p 3 D L U Z L V H F R P S D U L V R Q V K R Z H G between /p/ and /k/, and /p/ and /t/ (p L Q E R W K F D V H V + R Z H Y H not show such differences between singletons and geminates (p 3.05).

Figure 1. Distribution of log-transformed closure duration (left) and VOT (right) of singleton and geminate stops.

The mean and standard deviation of total C duration, CD and VOT are shown in Table 1 according to the place of articulation (POA). We could observe a shortening tendency of CD moving more posterior in POA, while VOT values described a reverse tendency. Both singletons and geminates showed this trend. The geminate CD was 157% of the singleton CD in bilabial, 177% in alveolar and 170% in velar stops.

	Total C duration		CD		VOT	
	S	G	S	G	S	G
[p]	97 (15)	139 (27)	76 (13)	119 (27)	21 (9)	20 (11)
[t]	91 (17)	137 (29)	65 (15)	115 (30)	26 (8)	22 (8)
[k]	98 (26)	147 (38)	63 (15)	107 (35)	35 (14)	40 (14)

Table 1. Mean and SD of total C duration, CD and VOT by POA.