The difference between Blends and Clipped Compounds

Hamans, C.

Published in:
Proceedings of IATL 34

Citation for published version (APA):
Proceedings of IATL 34

Papers from the 34th annual conference of the Israel Association for Theoretical Linguistics

Edited by Gabi Danon

Israel Association for Theoretical Linguistics
Ramat Gan, Israel
2020
# Table of Contents

Nasalization and Drawl in Central Yiddish  
*Guillaume Enguehard & Noam Faust* ................................................................. 1

Pre-nominal Adjectival Transparent Free Relatives  
*Alexander Grosu & Josef Bayer* ........................................................................ 17

The Difference between Blends and Clipped Compounds  
*Camiel Hamans* .................................................................................................. 31

Two Types of Object Incorporation in Uzbek  
*Zarina Levy-Forsythe & Olga Kagan* ................................................................. 45

A Note on Lexical Reciprocity in Brazilian Portuguese  
*Giada Palmieri, Renato Miguel Basso, Manuela Pinto, Yoad Winter & Joost Zwarts* .............. 67

Selectional Restrictions and the Meaning of Reciprocal Alternations  
*Yoad Winter* ...................................................................................................... 81
NASALIZATION AND DRAWL IN CENTRAL YIDDISH

GUILLAUME ENGUEHARD*1
NOAM FAUST**2
*University of Orléans, CNRS/LLL, **University of Paris 8, CNRS/SFL

1 Introduction

This paper deals with two phenomena in Central Yiddish (Jacobs 1990, 2005): nasalization and drawl. These phenomena are conditioned by syllabic weight and prepausal position, which at first sight implies a metrical process governed by moras and feet. Nevertheless, we will show that it is by no means necessary to use a prosodic hierarchy to account for them. Instead, direct, lateral relations between syllabic constituents suffice. An analysis is proposed within the autosegmental theory of strict CV (Lowenstamm 1996, Scheer 2004). Two lateral relations are assumed within this theory to account for the syllabic structure: licensing and government. We show that nasalization follows from licensing, and drawl follows from government.

The paper is structured as follows. In section 2, we provide the empirical data. Section 3 introduces the theoretical framework. Section 4 illustrates how the phenomena follow from the theory. The conclusion highlights two aspects of the account. First, seemingly metrical phenomena can be covered without reference to moras or feet. Second, the results carry implications for theories of consonantal strength.

1 Empirical data

All of the data in this paper come from work with a native speaker.3 The variety is that of the town of Plotsk, slightly to the north-west of Warsaw. As is typical of Central Yiddish, this variety exhibits five vowel phonemes /a,u,i,e,o/. Length is contrastive, e.g. /looz/ ‘louse’,

1 guillaume.enguehard@univ-orleans.fr
2 noam.faust@univ-paris8.fr
3 We thank our consultant Jean Spector for the long hours he spent with us exploring his native pronunciations. Other speaker who assisted in this study are (in alphabetical order) Dov Faust, Eliezer Niborski and Jean Lowenstamm.
There are also two diphthongs /oj/ and /aj/. A third diphthong [ej] is not contrastive with a long /ee/.

1.1 Nasalization

A consonant /n/ in word-final position is very often elided after diphthongs and long vowels (1a). Its underlying presence is nevertheless reflected in the nasalization of the vowel. The final /n/ reappears before vowel-initial suffixes, and the stem vowel is not nasalized (as a rule). Importantly, this alternation is not attested after short vowels (1b).

(1) Nasalization

<table>
<thead>
<tr>
<th></th>
<th>Unsuffixed</th>
<th>Suffixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsuffixed</td>
<td></td>
<td>Suffixed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fũũ</td>
<td>fuun-an^{PL}</td>
<td>‘flag’</td>
</tr>
<tr>
<td>diĩ</td>
<td>diin-an^{INF}</td>
<td>‘serve’</td>
</tr>
<tr>
<td>ūpjɔō</td>
<td>ūpjoon-an^{PL}</td>
<td>‘spy’</td>
</tr>
<tr>
<td>bãã</td>
<td>baan-an^{PL}</td>
<td>‘train’</td>
</tr>
<tr>
<td>tsej̃</td>
<td>tsejn-arl̩</td>
<td>‘ten’</td>
</tr>
<tr>
<td>vôj̃</td>
<td>vojn-an^{INF}</td>
<td>‘dwell’</td>
</tr>
<tr>
<td>vaj̃</td>
<td>vajn-an^{INF}</td>
<td>‘weep’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td></td>
<td>Suffixed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unsuffixed</td>
<td></td>
<td>Suffixed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nun</td>
<td>nun-an^{PL}</td>
<td>‘the letter י’</td>
</tr>
<tr>
<td>din</td>
<td>din-an^{PL}</td>
<td>‘law’</td>
</tr>
<tr>
<td>ton</td>
<td>ton-an^{PL}</td>
<td>‘ton’</td>
</tr>
<tr>
<td>man</td>
<td>man-an^{PL}</td>
<td>‘husband’</td>
</tr>
<tr>
<td>bren</td>
<td>bren-an^{INF}</td>
<td>‘burn’</td>
</tr>
</tbody>
</table>

Elision and nasalization also occur between a branching nucleus and a consonant-initial suffix, e.g. [fũũ-dl] ‘flag-DIM’. However, these phenomena are not attested after a nonbranching nucleus (e.g. [zĩnt] ‘sin’), and there are only few historical examples of them before a homomorphemic consonant (e.g. *frajnd > [fʁããt] ‘friend’). Thus, we can establish as a generalization that only root-final /n/ are subject to elision in Yiddish.

The pattern in (1) exhibits an interesting interplay between the realization of a final consonant and the length of a preceding vowel. Such interactions bring to mind the ban on trimoraic syllables discussed by Hayes (1989:291). If the ban on trimoraic syllables is admitted in Central Yiddish, nasalization results from a repair mechanism that delinks the final /n/ when it follows a bimoraic nucleus. However, one might raise the following questions: i. why can bimoraic nuclei be followed by final consonants other than /n/, and ii. why is the problem solved by nasalisation, as opposed to Closed Syllable Shortening? In this paper, we adopt a theory that does not admit moras at all, but proposes a solution to these two issues.

---

4 Short vowels may be realized laxer than long vowels, e.g. [loos] ‘louse’, [loz] ‘let’. In our transcriptions, we ignore this non contrastive phonetic issue.

5 ‘tenner (for instance for a ten dollar bill)’
1.2 Drawl

Drawl is the process whereby a vowel is broken in two, resulting in hiatus: /vuus/ ‘what’ is pronounced [vúuəs]. Central Yiddish exhibits drawl in stressed, long vowels before a pre-pausal consonant. The vowels that undergo drawl are /uu,oo/ before the coronal consonants /t,d,s,z,l/ (2a), and /uu,oo,ii,ee/ before the uvular consonants /ʁ,χ/ (2b).\(^6\)

(2) Drawl

<table>
<thead>
<tr>
<th></th>
<th>#</th>
<th># ##</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>/_s/</td>
<td>vuus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>aroos</td>
</tr>
<tr>
<td></td>
<td>/_z/</td>
<td>nuus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>loos</td>
</tr>
<tr>
<td></td>
<td>/_t/</td>
<td>ſtuut</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hoot</td>
</tr>
<tr>
<td></td>
<td>/_t/</td>
<td>buut</td>
</tr>
<tr>
<td></td>
<td></td>
<td>moot</td>
</tr>
<tr>
<td></td>
<td>/_l/</td>
<td>muul</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fool</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>#</th>
<th># ##</th>
</tr>
</thead>
<tbody>
<tr>
<td>b.</td>
<td>/_χ/</td>
<td>biiχ</td>
</tr>
<tr>
<td></td>
<td></td>
<td>nuuχ</td>
</tr>
<tr>
<td></td>
<td>/_ʁ/</td>
<td>fiiʁ</td>
</tr>
<tr>
<td></td>
<td></td>
<td>puuʁ</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dooʁ</td>
</tr>
<tr>
<td></td>
<td></td>
<td>veεʁ</td>
</tr>
</tbody>
</table>

All final stressed vowels in Central Yiddish are long. They never undergo drawl (3a). Short vowels never undergo drawl (3b).

(3) No drawl

<table>
<thead>
<tr>
<th></th>
<th>#</th>
<th># ##</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>bluu</td>
<td>bluu ‘blue’</td>
</tr>
<tr>
<td></td>
<td>fʁoo</td>
<td>fʁoo ‘wife’</td>
</tr>
<tr>
<td></td>
<td>kii</td>
<td>kii ‘cow’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>#</th>
<th># ##</th>
</tr>
</thead>
<tbody>
<tr>
<td>b.</td>
<td>nul</td>
<td>nul ‘zero’</td>
</tr>
<tr>
<td></td>
<td>los</td>
<td>los ‘let’</td>
</tr>
<tr>
<td></td>
<td>ftrʁiχ</td>
<td>ftrʁiχ ‘line’</td>
</tr>
</tbody>
</table>

The two effects, nasalization and drawl, are combined in /uu,oo/ before final /n/ (4). This is surprising: the nasalized vowel is final, and should behave like the vowels in (3a). Instead, it behaves like the vowels in (2a), as if it were followed by a consonant.

---

\(^6\) Jacobs (2005:96) points out that the realization of drawl can regionally vary between V:\, VG\ (G = glide) and Va (e.g. /buud\ → [buuəd] ~ [buwəd] ~ [buəd]). Here we adopt Jacobs’ (2005) transcription [V:\].

\(^7\) We found no case of long vowels occurring before an affricate [ʦ].
Like nasalization, drawl also concerns super-heavy (C)VVC syllables. The reference to syllabic weight suggests that metrical structure is relevant: a foot, based on syllables and segments that project moras or not. However, such structure is not sufficient to explain why i. schwa occurs after a long vowel only in prepausal position, and ii. schwa occurs only before coronal and uvular consonants. We claim that a flat representation of the word structure with lateral relations between syllabic constituents can account for these facts. In the next section, we present the basic tenets of the relevant theoretical framework, Strict CV.

2 Theory

2.1 Government

Like all autosegmental theories, Strict CV (Lowenstamm 1996, Scheer 2004) distinguishes between segments and skeletal positions, and places them on different tiers. Unlike in other autosegmental theories, in Strict CV there is only one skeletal unit: the CV unit. Thus, the skeletal tier of all items begins with a C-slot and ends with a V-slot. Accordingly, segmentally consonant-final words are skeletally V-final (5).

(5) Final empty nucleus: [kij] ‘kiss’

C V C V
k i j

Empty V-slots are not freely distributed in representations. In language like Central Yiddish, which allow consonant-final words, final V-slots are allowed by assumption. They are called Final Empty Nuclei, henceforth FEN.

Assuming CV to be the only segmental unit leads to medial Empty V-slot, too. This is the case in words with phonetically adjacent consonants (6). In order to avoid larger groups of consonants, Kaye et al. (1989) propose that empty V-slots can only remain empty under certain conditions. When these conditions are not met, epenthesis may surface. Whether this will occur depends on the environment. If the next V-slot is contentful, then an empty V-slot can remain silent (6a). This relation is called “government”: contentful Vₙ governs empty

---

8 Drawl is less audible after [oo]. While the speaker we worked with was hesitant about drawl in these words, we have encountered occurrences of it in other speakers.
V_{n-1}, thereby inhibiting its realization. However, if V_n is itself empty, an empty V_{n-1} cannot remain silent and either V_n or V_{n-1} will be realized through epenthesis (6b).^9

(6)  
\begin{align*}
a. & \text{[gazlən] ‘robber’} \\
& \text{C V C V C V C V} \\
& g a z l a n \\

b. & \text{[majdələ] ‘young lady (dim.)’} \\
& \text{C V C V C V C V} \\
& m a j də lə \\
\end{align*}

We’ve seen the representation of i. words ending in a consonant and ii. words with adjacent consonants. The two conditions can be combined: a word can end in adjacent consonants. In that case, as shown in (7), the last two V-slots of the skeleton are empty. The ultimate or the penultimate V-slot is therefore expected to be realized. Indeed, many languages exhibit epenthesis in final underlying cluster. However, many other languages, and Yiddish among them, do admit final clusters. Within Strict CV, this means that the FEN, although empty, can govern the preceding empty nucleus. The ability of the FEN to govern may be regarded as following from the property that it shares with contentful nuclei: it is not governed.

(7)  
\[\text{[zamʃ] ‘suede’}\]

\begin{align*}
& \text{C V C V C V} \\
& z a m ŵ \\
\end{align*}

2.2 Licensing

Length in Strict CV is straightforwardly represented as one-to-many association between a segment and two slots. Thus, the word [siidə] ‘feast’ is represented as in (8a). Many languages, including Yiddish, exhibit a ban on long vowels in closed syllables. This restriction is one possible manifestation of the ban on trimoraic syllables mentioned earlier. Rather than referring to the prosodic structure of the word, Lowenstamm (1996) links this effect, too, to the status of the following nucleus. The length of a vowel is licensed by a following contentful nucleus. Scheer (2004) argues that this lateral relation must be distinct from government: government inhibits realization, whereas in this case the realization is assisted. He dubs this relation “licensing”. Closed syllable shortening thus follows from the lack of licensing from the following nucleus. This is illustrated for the impossible word *[míiʃpat] in (8b).

---

^9 The case in (6b) is taken from Faust (2018). He analyzes this form as a “double diminutive”, involving the addition of diminutive suffix /l/ not once but twice. This scenario yields three consecutive empty nuclei. The final [ə] is inserted in order to avoid haplology in [ləl]. The medial [ə] is inserted because its position is not governed. Cases of Yiddish epenthesis are rare and conditioned by multiple factors. We do not intend to provide a complete explanation of this phenomenon here.
As in many languages with closed syllable shortening, final consonants in Central Yiddish do not count as closing the syllable. Thus, as was already shown in (3) above, long vowels may precede a final consonant. This is in fact predicted by the model. We've seen that FENs behave like contentful nuclei for the purpose of government, it is therefore not surprising that they can also license (9).

Finally, just as the FEN can be empty without being governed, it can also host vowel spreading without being licensed. This aspect is represented by the long final vowel of [fʁoo] in (10).

We have introduced the notions of government and licensing. We will now see that government accounts for drawl, and licensing for nasalization.

3 Analysis

3.1 The representation of /n/

Before addressing the issue of nasalization, it is important to understand why only /n/ drops in the final position after a long vowel. This consonant has a tendency to be homorganic both cross-linguistically and in Yiddish (Jacobs 1995:113-114). This tendency reveals the absence of an underlying place feature in /n/, which is later acquired as a result of assimilation or default insertion. But what remains in /n/ if we remove the place of articulation? Voicing is not contrastive in Yiddish nasals, and so, when stripped of its place, /n/ is essentially made of a lone mode feature: nasality. Interestingly, the fewer features the consonant bears, the more
likely it is to i. be weak (Harris 1990), and ii. behave like a vowel (Pöchtrager 2001). These two properties are observed for /n/ in Central Yiddish.

First, Central Yiddish nasals have the ability to license a long nucleus as vowels do. Consider the representation of [liign] in (11a). The empty nucleus to the left of /n/ being governed, it should not be able to license the preceding nucleus. Therefore, one does not expect to find a long vowel in the first syllable. But, if we accept Pöchtrager’s (2001) proposal, the final nasal is able to spread to the nucleus to its left (11b). This latter being filled, it does not need to be governed and can therefore license the long vowel of the first syllable.

(11) a. *[lign] ‘lie’  
  C V C V C V C V  
  l i g n  
  G

b. [liign] ‘lie’  
  C V C V C V C V  
  l i g n  
  L

Thus, we assume the parameter in (12) to be active in Central Yiddish.

(12) |N| can fill an empty nucleus.

Secondly, Central Yiddish /n/ is weak enough to drop. Following Scheer (2004) and Faust & Torres-Tamarit (2017), we will assume that /n/ is a weak consonant that cannot be linked to its position unless it is licensed by a following vowel. However, it is not clear why unlicensed /n/ only drops in the final position: compare [diĩ] and [diĩ-n-ən] ‘serve’. In the second case, /n/ is unlicensed but it does not drop. We have no formal motivation for this fact. However, it can be linked to another phenomenon of the same language, namely final devoicing. In Element Theory (Kaye et al. 1985, 1989; Nasukawa 1997; Backley 2011), nasality and voice are both represented by an element |N|, which is responsible for the decrease in intensity in the high frequencies of the spectrum. Some Central Yiddish dialects, including the one under discussion, exhibit the final devoicing of obstruents. A representation of this final devoicing in Element Theory is proposed in (13). Voiced obstruents maintain their element |N| in medial position (13a) but they lose it in final position (13b). The reason that final devoicing does not involve nasalization of the vowel in the final syllable is the difference between the |N| of voicing and that of nasality: the first is a head element while the second is an operator (Backley 2011:145-157).

(13) a. [huuz-ənə] ‘rabbit-like’  
  C V C V C V C V  
  h u h e n e  
  A
  N

b. [huus] ‘rabbit’  
  C V C V C V  
  h u h  
  A
  N

Thus, n-dropping and final devoicing can be unified with the language-specific rule in (14).
(14) Final |N| is afloat (unassociated) underlyingly

Based on the specific generalizations in (12) and (14) concerning the representation of nasals, we will show that the relation between nasalization and drawl follows the principles of Government Phonology.

3.2 Nasalization and licensing

We discuss below cases like [din] (i.e. no final /n/-dropping after a short vowel), [dīī] (i.e. final n-dropping after a long vowel) and [diinən] (i.e. no medial n-dropping). As shown in (9) above, a FEN is a legitimate licensor in Yiddish. Thus, in (15), a final /n/ after a short vowel is licensed and can be linked to the C-slot. This is also the case of the syllabic nasal in (11).

(15) [din] ‘law’

\[
\begin{array}{c}
\text{L} \\
\text{C V C V} \\
\text{d i N}
\end{array}
\]

We also saw that licensing from the FEN is required in order to maintain a long vowel before final consonants. This leads to the conflict in (16): in a sequence /VVn#, both the nasal and the long vowel before it require licensing from the FEN. Only one of the two can be licensed. It seems that the vowel is prioritized in this situation. As a consequence, the nasal may not associate to its position. Instead, it remains afloat and is realized as nasalization on the preceding vowel.

(16) [dīī] ‘serve’

\[
\begin{array}{c}
\text{L} \\
\text{C V C V C V} \\
\text{d i N}
\end{array}
\]

When the same noun is suffixed, as in (17), the nasal is not final. It therefore does not need to be floating, and nasalization does not occur systematically.

(17) [diin-ən] ‘serve-INF’

\[
\begin{array}{c}
\text{L} \\
\text{C V C V C V} \\
\text{d i n a N}
\end{array}
\]
Licensing underlies nasalization and its distribution in Central Yiddish. It motivates the absorption of a final nasal into a preceding long vowel, but its retention after short vowels.

### 3.3 Drawl and government

In this subsection, we discuss cases such as: [vuus] (i.e. drawl before a prepausal coronal consonant), [bluu] (i.e. no drawl in final position), [nul] (i.e. no drawl after a short vowel), [fuũa] (i.e. the correlation between drawl and nasalization) and [fiiaʁ] (drawl before non-coronal consonants). In order to understand drawl, one must first understand pause. In line with Lowenstamm (1999) and Scheer (2000), Pagliano (2003) argues that phonological effects at the utterance level must follow from phonological representations (rather than diacritics signaling frontiers, such as the prosodic hierarchy). She argues that edges and pauses might be phonologized through the addition of skeletal units. Accordingly, we will hypothesize here that pause is exponed by the only skeletal unit, the CV unit.

The addition of the CV unit of the pause has the following effect. The nucleus after the final consonant, which in context is the FEN (18a), is no longer final in (18b). It becomes a regular empty nucleus. Non-final empty nuclei need to be governed. When governed, they may neither license nor govern. Thus, being unlicensed, the second nucleus of the long vowel in (18b) cannot be occupied by it; nor can it remain empty, because it is ungoverned. It is neither engaged nor inhibited, and must therefore be realized. This is drawl.

\[(18)\] 

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[C V C V C V]</td>
<td>[C V C V C V [C V]]</td>
</tr>
<tr>
<td>v u s</td>
<td>v u a s</td>
</tr>
</tbody>
</table>

Note that the derived form in (18b) is [vuas], not [vuus]. As reported in fn. 6, Jacobs (2005) mentions regional variation between Vːa, VGa and Və. This variation can be captured by the theory. After the schwa is inserted, the preceding vowel can spread to the C-slot on its right to avoid hiatus (19). Provided that the distinction between a high vowel and its glide equivalent is not unequivocal, this spreading can be interpreted as in [vuwas] or [vuas]. We will not indicate this spreading in the following representations.

\[(19)\] 

[vuwas]~[vuus] ‘what’

| [C V C V C V [C V]] |
| v u a s |

The representation of drawl in (18) explains why it does not occur in final position or within a short vowel. We saw that, when it is in final position, a long vowel does not need to be licensed (20a). In a prepausal context, its second position is not an FEN anymore and it does
Guillaume Enguehard & Noam Faust

need to be licensed in order to host a long vowel (20b). This is made possible by the FEN on its right which, as we saw in the previous section, is a legitimate licensor.

(20) a. [bluu] ‘blue’ b. [bluu] ‘blue’

After a short vowel, the only empty nucleus is the final one (21a). In prepausal context, this nucleus is governed by the FEN and nothing happens (21b).

(21) a. [nul] ‘zero’ b. [nul] ‘zero’

We already saw that the loss of final /n/ after a long vowel and the nasalization of that vowel are due to the fact that the FEN needs to licence the long vowel to the detriment of the nasal consonant (22a). In prepausal context (22b), neither the nasal consonant nor the long vowel can be licensed or governed by the nucleus on their right because the latter is governed by the FEN of the pause. Therefore, none of them can be realized. The second position of the long vowel is filled by an epenthetic vowel and the nasal element is absorbed by the nucleus. We further assume that nasalization spreads through the hiatus and reaches the lexical vowel as well.

(22) a. [fũũ] ‘flag’ b. [fũũə̃] ‘flag’

At this stage of the analysis, it seems impossible for a final nasal to be realized as a consonant in prepausal context, even if it is preceded by a short vowel (e.g. /din##/ ‘law’). Since the nucleus on the right of /n/ is governed by the FEN, it may not license the preceding nasal (23a). However, we saw that |N| can spread to a V position and license a preceding position. This is represented in (23b): |N| can spread to the nucleus position on its right and license its own onset slot. This situation is not possible in (22b) since the nasal would necessarily license the preceding long vowel to its own detriment, thus being unable to occupy its own position (licensing may not have two targets simultaneously).

---

10 See Scheer (2000) for a similar analyses for long vowels in languages without Closed Syllable Shortening.
Nasalization and drawl in Central Yiddish

Up to now, we discussed drawl without regards to the type of consonant which follows the underlying long vowel. But the distribution of drawl is slightly different before coronal on the one hand and before /ʁ,χ/ on the other. Before coronal consonants, only /u,o/ are impacted by the occurrence of a schwa in prepausal context. Before /ʁ,χ/, /i,u,o/ are impacted by this phenomenon. The questions are: i. why does drawl depend on the quality of the following consonant, and ii. why does drawl depend on the quality of the preceding vowel? We hypothesize that this is due to the internal composition of the different segments. An inventory of Yiddish consonants and vowels in terms of their Elemental constitution is provided in (24). The two classes of consonants which trigger drawl contain an element |A| (operator in coronals and head in uvulars). Interestingly, the schwa of the drawl also contains |A|. We conclude that the quality of the epenthetic vowel observed in drawl partly depends on the quality of the following consonant.

When there is no observed drawl, it implies that the quality of the epenthetic vowel is conditioned by the preceding vowel. In other terms, there is a competition between the place elements of the preceding vowel and those of the following consonant to be absorbed by the epenthetic vowel. The case of [vuua] in (25a) suggests that |A| takes precedence over |U|. The case of [ziis] in (25b) suggests that |I| takes precedence over |A|, and the case of [biiaχ] in (25c) suggests that |A| takes precedence over |I|.

To sum up, the hierarchy between elements follows the pattern in (26).
This pattern leads to the predictions in (27). The black boxes represent long vowels followed by a palatal consonant. Such forms are not attested in Central Yiddish and the prediction cannot be confirmed or denied. White boxes represent cases where the element of the consonant takes precedence over the elements of the preceding vowel. This concerns /ii,ee,oo,uu/ before a uvular consonant and /oo uu/ before a coronal consonant. Finally, grey boxes represent cases where the elements of the consonant do not take precedence over the elements of the preceding vowel. These forms have no observable drawl because the epenthetic vowel copies the quality of the preceding vowel.

<table>
<thead>
<tr>
<th>The final coda contains...</th>
<th>The vowel contains...</th>
</tr>
</thead>
<tbody>
<tr>
<td>[A] faaʁ ‘fire’</td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>ίiis ‘sweat’</td>
</tr>
<tr>
<td></td>
<td>[A] vaab ‘wife’</td>
</tr>
</tbody>
</table>

To sum up, the syllabic conditioning of drawl, i.e. the fact that it affects long vowels in a prepausal context, is a correlate of lateral relations. Its segmental conditioning, i.e. its limitation to back, round vowels before coronals and high vowels before uvulars, results from the competition between elements to fill the ungoverned position.

**Conclusion**

We have discussed nasalization and drawl in Central Yiddish. The two phenomena involve the same configuration: a long, closed syllable at the edge of the word. Elsewhere in the literature (e.g. Hayes 1995), such syllables are referred to as "super heavy". Their heaviness is expressed in the number of weight units that they involve. In the framework adopted here, there are no weight scales, no weight units ("moras"), and no feet; only lateral relations between syllabic constituents. Using these latteral relations, we've accounted for the seemingly metrical phenomena at hand. We've shown how drawl follows from government and nasalization from licensing. No appeal to metrics-specific vocabulary was necessary. In that sense, our account shows that phenomena usually analyzed using metrical mechanisms do not necessarily depend on such arborescent representations. Instead, they can be captured using a flat representation of phonology. This conclusion joins previous work in challenging such an arborescent conception of metrics (Chierchia 1986, Larsen 1998, Scheer & Szigetvári 2005, Ulsfsbjorninn 2014, Enguehard 2016, i.a.).
The analysis also carries a consequence for theories of consonantal strength. Such theories generally describe lenition and fortition processes as conditioned by three types of syllabic positions: i. the non-intervocalic onset, ii. the intervocalic onset, and iii. the coda. The implementation of this assumption in Strict CV gave birth to Coda Mirror Theory (Ségéral & Scheer 2001). Based on the idea that Government and Licensing condition the realization of segments, Coda Mirror Theory makes the prediction in (28). The non-intervocalic onset is strong because it is not inhibited by Government, the intervocalic onset is weak because it is inhibited by Government, and the coda is also weak because it is not assisted by licensing. Ségéral & Scheer notice that their theory predicts a fourth logical possibility which they consider to be unattested, namely an inhibited and unassisted position they call “nightmare position”.

<table>
<thead>
<tr>
<th>Lateral relations</th>
<th>Syllabic position</th>
<th>Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>-G+L</td>
<td>non-intervocalic onset</td>
<td>strong</td>
</tr>
<tr>
<td>+G+L</td>
<td>intervocalic onset</td>
<td>weak</td>
</tr>
<tr>
<td>-G-L</td>
<td>coda</td>
<td>weak</td>
</tr>
<tr>
<td>+G-L</td>
<td>nightmare position</td>
<td>weak</td>
</tr>
</tbody>
</table>

We saw that in Central Yiddish, the realization of final /n/ depends on the length of the preceding vowel. Indeed, in our analysis the consonant competes with the preceding vowel for the licensing emanating from the following nucleus. The consonant loses, licensing goes to the long vowel. Assuming that the active FEN must dispense its governing force, the final /n/ is precisely in the nightmare position (29). Our account thus confirms the prediction of Coda Mirror Theory in that it identifies a consonant in this previously unattested configuration.

Thus, later improvements of Coda Mirror Theory striving to get rid of the nightmare position (e.g. Scheer & Ziková 2010) should take into account the effect of a preceding vowel on the consonant. On the bases of the Yiddish facts, one might conclude that a consonant in the nightmare position, i.e. occurring after a long vowel, can be weaker than a consonant occurring after a short vowel. Time will tell if this finding is further corroborated by investigations into other lenition phenomena.

References


1 Introduction

This paper is a follow-up to Grosu (2016), and focusses on one aspect of the theory of Transparent Free Relatives (TFRs) that was analyzed in insufficient detail in Grosu (2016) (as well as at the oral presentation based on that paper, which was made by Alexander Grosu at IATL 34). The contributory thrust of the paper concerns adjectival TFRs in pre-nominal position. Much as was done in Grosu (2016) with respect to nominal entity-denoting TFRs, this paper will examine the merits and/or demerits of two competing theories of TFRs insofar as the analysis of pre-nominal adjectival TFRs is concerned.

The paper is structured as follows: Section 2 presents the gist of the proposals in Grosu (2016) that constitute minimally necessary background for ensuring that this paper is self-contained. Section 3 discusses pre-nominal adjectival TFRs. Section 4 is a summary of results.

2 Background Information

TFRs have been the object of numerous earlier studies, an incomplete list being Nakau (1971), Kajita (1977), McCawley (1998), Wilder (1998), van Riemsdijk (1998, 2000, 2001, 2006a,b, 2012, 2017), Grosu (2003, Part II, 2010, 2014, 2016), Schelfhout, Coppen and Oostdijk (2004), den Dikken (2005), van de Velde (2011), Smet and van de Velde (2013), and Yoo (2008). In this paper, we will be concerned with a special type of TFR, in particular, the one described in the title of the paper, and – in the spirit of Grosu (2016) – we will confine ourselves to considering the implications of the relevant empirical facts for two of the existing theories of TFRs: the one
defended in Grosu (2016), and the one argued for by Henk van Riemsdijk in the various studies mentioned in this paragraph.

As indicated in the introduction, this section summarizes the principal claims made in Grosu (2016) concerning the merits and/or demerits of the two theories with respect to nominal entity-denoting data. As also noted in the introduction, this section presents what we view as minimally necessary background for reading section 3 without consulting the earlier literature. For completeness, we note that readers might nonetheless benefit from familiarizing themselves with the pertinent earlier literature.

Illustrations of TFRs, which provide a basis for indicating some of their distinguishing properties, are presented in (1).

(1) a. He is eating [what my grandfather thinks [t is a pork chop]].
   b. He is eating [what I would call [t a huge steak]].

In contrast to (non-transparent) free relatives (FRs), TFRs are necessarily introduced by what (and its cross-linguistic counterparts) 1, the trace of what is necessarily in the subject position of a copular structure or small clause (as in (1a) and (1b) respectively), and the relative clause must include an explicit or implicit intensional operator (italicized in (1)), without which it is infelicitous. Of course, FRs may also exhibit these properties, potentially giving rise to ambiguity, but they do not in general have to exhibit them (demonstration omitted).

One sub-element of TFRs that the two theories under consideration analyze very differently is the non-subject of the copular construction or small clause, which appears in boldface in (1), and to which we will refer with the pre-theoretical term 'pivot'. Thus, Grosu (2016) assigns to TFRs the same gross syntactic structure as to FRs, in particular, the one schematically indicated in (2). On this view, the TFRs in (1) are complex DPs headed by a null Det2, and the pivot (in boldface) is just what it seems to be, namely, the non-subject of a copular sentence/small clause.

(2) \[ [\text{DP 0}_{\text{DET}} [\text{CP what} \ldots [\text{BE YP} \ldots \ldots]]] \]

Van Riemsdijk assigns to TFRs and FRs different syntactic structures (the structure assigned to FRs will not concern us in this paper). For TFRs, he assumes a multi-dimensional framework, in which different clauses may lie in different planes. For TFRs like those in (1), the matrix and the relative clause lie in two distinct planes, as suggested by the schema in (3), where the two clauses do not form a constituent. The clauses are only related by an operation of 'grafting', whose effect, in TFRs, is to re-merge an element of the relative clause, in particular, the pivot, with some sub-element of the matrix clause, with the result that the pivot constitutes 'shared structure.' In (1), the matrix part of the pivot is the direct object of the matrix verb, and plays a role, mutatis mutandum, analogous to that of a relative-external 'phrasal head' in a bi-dimensional framework.

(3) \[ [\text{Matrix Clause} \ldots [\text{YP}_k \ldots \ldots]] \]
   \[ [\text{CP what}_i \ldots [\text{BE YP}_i \ldots \ldots]] \]

---

1 See, however, Schütze & Stockwell (2019) on TFRs introduced by who in English.
2 While the TFRs in (1) are nominal, TFRs of other categories also exist, in particular, adjectival and adverbial ones. In the latter two cases, the null head of the XP is suitably different.
In our view, the fundamental difference between the two theories lies in their views on the 'headedness' of TFRs. While Grosu views TFRs as headed by a null category, van Riemsdijk views them as 'quasi-headed' by the pivot. The bi- versus multi-dimensional distinction between the two theories follows from the fact that in a bi-dimensional framework, it is unclear how to represent the pivot as a CP-external head when it occupies a string-medial position within the relative, as in the English and German data in (4).

(4) a. I just saw [what might well be taken for a meteor by my neighbours].
   b. Ich werde mir kaufen, [was du als einen passenden Wagen bezeichnen würdest].
      I will buy what you as a suitable car characterize would
      ‘I will buy myself what you would describe as a suitable car.’

The two approaches are also driven by different considerations. For van Riemsdijk, the primary motivation for viewing the pivot as an element shared by the relative and the matrix was provided by a number of syntactic effects that are typically associated with CP-external heads of complex XPs, for example, the fact that the categorial properties of the TFR necessarily match those of the pivot, and the fact that the pivot agrees in syntactic number with the matrix verb under certain circumstances. An additional motivation was the belief that the pivot is not only syntactically present in the matrix, but also interpreted there. For example, (1b) was viewed as having the essential import of (5a), where the TFR is paraphrased by a parenthetical with 'hedging' import. Schelfhout et al (2004) in fact took this view one step further and proposed that the TFR is not only paraphraseable by a parenthetical sentence, but is in fact a parenthetical modifier of the pivot, and furthermore claimed that data like (1) are necessarily pronounced with slight intonational breaks in the positions indicated by commas in (5b).

(5) a. He is eating a huge steak, at least, this is what I would call it.
   b. He is eating, what I would call, a huge steak.

For Grosu (2016), the primary motivation for viewing the pivot as present in the relative clause only (at all levels of representation) was semantic. It was pointed out in that study that a parenthetical paraphrase of the kind illustrated in (5a) is completely impossible in some cases, and is also inappropriate in numerous additional cases, albeit more subtly. An example of the former kind is provided by a slight modification of (1a), shown in (6a), which clearly does not have the import of the self-contradictory paraphrase in (6b). An example of the latter kind is (7a), which differs in meaning from the paraphrase in (7b) in that uttering it in no way commits the speaker to the belief that ghosts exist; in contrast, uttering (7b) does commit the speaker to that belief. Note that in (7b), the speaker begins by asserting that 'he' saw a ghost, and then hedges on whether the entity at issue was indeed a ghost, but cannot reject the assumption that ghosts exist without contradicting himself/herself. We suggest that the impression of synonymy between (1b) and (5a) (1b) stems from the fact that the speaker not only asserts the main clause, but is also responsible for the content of the subordinate clause (as its grammatical subject), and may thus be viewed as parenthetically asserting it. – See Grosu (2016 pp. 1254-5 for more detailed illustration and discussion of these matters.
(6) a. He is eating [what my grandfather incorrectly thinks [t is a pork chop]].
   b. #He is eating a pork chop, as my grandfather incorrectly thinks.

(7) a. He saw [what he believes was a ghost].
   b. He saw a ghost, at least, this is what he believes.

As just suggested, the impression that the speaker of (1b) asserts that 'he' is eating a huge steak arguably arises from the fact that the speaker is also the person who calls the eaten object that way. It suffices, however, to use a different expression as subject of the subordinate verb to eliminate that impression. Thus, the example in (8) is in no way contradictory.

(8) He is eating what Mary would probably call a huge steak, but I disagree, that piece of meat is not particularly large, and moreover is very probably not even a steak.

We have the following comment about the kind of parenthetical analysis proposed by Shelfhout et al (2004), which, we note, was not discussed in detail in Grosu (2016): Insofar as the kind of intonation in (5b) is concerned, we agree that it is possible, but it is certainly not required (according to our intuitions, and those of numerous speakers of various languages that we have consulted). Furthermore, we do not see how data like (4) could even be uttered with a suitable parenthetical intonation (Shelfhout et al do not discuss this point).

As far as interpretation is concerned, we draw attention to the fact that the intonational contour at issue makes possible at least two distinct construals, which require radically different analyses (see below). Thus, one construal of (5b) (presumably, the one intended by Shelfhout et al), is essentially equivalent to (5a), with the difference that the hedge is expressed before the assertion of the pivot. Another construal is that the speaker wishes to create suspense before uttering the pivot. It is important to note that if data like (6a) are modified to fit the pattern in (5b), as in (9a), the former kind of construal is impossible, being self-contradictory (just like (6b)), but the latter type of construal is possible, as brought out more explicitly by (9b).

(9) a. He is eating, what my grandfather incorrectly thinks is, a pork chop.
   b. He is eating, what my grandfather incorrectly thinks is … a rat, of all things!
   Incredible, what these old people can imagine!

Grosu (2016) rejected the view of TFRs as parenthetical hedges, and proposed instead that their raison d'être is to characterize two potentially distinct guises/counterparts of something, which exist at distinct sets of worlds/indices, the guise denoted by the TFR being indeterminate/unspecified at the indices of the matrix. In data like (1a), the counterparts are defined in (potentially) distinct ways in the speaker's and the grandfather's belief-worlds, while in data like (1b), the distinct sets of indices are the belief-worlds of the speaker and those of implied individuals who might disagree with the way in which the speaker characterizes what is eaten. The need for two at least potentially distinct guises is brought out by the fact that when no distinct indices are implied, the result is infelicitous, as in (10a), presumably due to triviality. It suffices, however, to imply the existence of distinct indices by means of focus for infelicity to disappear, as in (10b).
Grosu (2016) further proposes that the indeterminate/unspecified status of the guise denoted by the TFR is reflected in the following fundamental difference between FRs and TFRs: While FRs are necessarily construed as definite (a view widely defended in the literature), TFRs are invariably construed as (nonspecifically) indefinite. This distinction is detectable in ambiguous constructions, the context sometimes favouring one construal over the other. For example, if the speaker of (1a) cannot see what 'he' is eating, but was just told by his/her grandfather that it is a pork chop, this example is naturally construable as including a TFR, and is paraphraseable by (11a). On the other hand, if the speaker and the grandfather had the opportunity to examine the meal prior to its being eaten by 'him', and if the grandfather, but not necessarily the speaker, thought it was a pork chop, the example is naturally construable as including an FR, and is paraphraseable by (11b). Correspondingly, the null Det in (2) is construed as a definite operator in FRs and as an indefinite/existential one in TFRs.

Thus, Grosu's view that the pivot belongs only in the subordinate clause was primarily motivated by a demonstration that it is exclusively interpreted there, no part of it being construed in the matrix (for illustration and discussion, see Grosu 2016, section 3.1). An additional motivation for adopting this view was the existence of certain syntactic facts that conflicted with the assumption that the pivot is syntactically present in the matrix (for details, see Grosu 2003, sections 5.4 and 5.5, or Grosu 2016, pp. 1252-3). With respect to the syntactic effects that are consistent with the view that the pivot is present in the matrix, and may in fact be used as supporting arguments for that view, Grosu (2016) proposed that what in TFRs, unlike what in FRs and interrogatives, is unspecified for category, syntactic number and non-human status. He further proposed that this under-specification enables what to 'inherit' certain properties from the pivot, regardless of whether the copular structure or small clause is construed equationally or predicatively, and that this state of affairs gives rise to what may be called a 'transparency channel', through which properties of the pivot may be 'conveyed' to the TFR, and conversely.

We conclude this section with two important remarks.

The first is that from the perspective of Grosu's (2016) characterization of TFRs, parenthetical constructions like (5b) are not TFRs at all on the construal that involves the interpretation of the 'pivot' at matrix indices. On the construal where the uttering of the pivot is delayed for suspension effects, the pivot is interpreted at the indices of the relative clause (as brought out by the fact that data like (9a-b) are possible), and the construction is a bona fide TFR.

---

3 This is an oversimplification of the semantic analysis in Grosu (2016), which, however, will do for present purposes. In Grosu (2016), it is proposed that the relative CP denotes a set of individual concepts (of type <s,e>), and that the value of the complex DP at matrix indices may be indefinite/indeterminate, even if the Det that takes CP as argument is definite, with the result that the intensional object denoted by the complex DP is unique. In such a case, indeterminacy results from the possibility that the values of that intensional object may vary across the speaker's belief-worlds.

4 The latter possibility is needed for handling pre-nominal adjectival TFRs; see section 3.
The second is that from the perspective of van Riemsdijk's theory, data like (6)-(7) can only be handled, as far as we can see, by admitting that the pivot, even if syntactically present in the matrix, cannot be interpreted there, at least, not in general (see Grosu 2016, section 6, for elaboration of this point).

3 Pre-Nominal Adjectival TFRs

We now turn to a consideration of adjectival TFRs, which, like APs in general, may occur either in predicative or in ad-nominal position, as in (12)-(13) respectively.

(12) This story is [what many people might consider highly intriguing].


We note that some speakers view data like (13a), in which the TFR immediately follows a determiner, as somewhat marginal, but are more ready to accept data like (13b), where the TFR is coordinated with a preceding lexical adjective.

A fact of some importance, which was not explicitly pointed out by Grosu (2016), is that adjectival TFRs yield essentially the same kind of semantic support for Grosu's theory as the data that were brought up in section 2. Thus, adjectival TFRs are infelicitous if the relative includes no obvious explicit or implicit intensional operator (see the (a) sub-cases of (14)-(15)), are compatible with situations in which the speaker subscribes to the content of the pivot (see the (b) sub-cases of (14)-(15)), and are also compatible with situations in which the speaker decidedly rejects the content of the pivot (see the (c) sub-cases of (14)-(15)). Furthermore, data like (15c) are not an idiosyncratic property of English; they are also found, for example, in French and Romanian, as illustrated by the (b) sub-cases of (16)-(17).

(14) a. #This story is [what is interesting].
    b. This story is [what I might call interesting].
    c. This story is [what no one in his right mind would ever call interesting].

(15) a. #Bill proposed a [(new and) what is interesting] solution.
    c. Bill proposed a [(crazy and) what no one would ever call interesting] solution.

(16) a. Jean a fait une (nouvelle et) ce que j'appellerais très intéressante proposition.
    Jean has made a new and Dem that I-would call very interesting proposal
    'Jean has made a new and what I would call very interesting proposal.'
    b. Jean a fait une (stupide et) ce que personne n'appellerait intéressante proposition.
    Jean has made a stupid and Dem that nobody Neg-would-call interesting    proposal
    'Jean has made a stupid and what no one would call interesting proposal.'
(17) a. Ion a prezentat o (nouă și) ceeace aș numi extrem de interesantă soluție.
   "Ion has presented a new and what I would call extremely interesting solution."

   b. Ion a prezentat o (veche și) ceeace nimeni n-ar considera interesantă soluție.
   "Ion has presented an old and what nobody would consider interesting solution."

The importance of data like (14)-(17) lies in the fact that they broaden the empirical data-base of Grosu's thesis, which holds that the pivot needs to be interpreted in the relative. In particular, they show that the pivot must be so interpreted not only in nominal, but in adjectival TFRs as well.

Having established that much, we now propose to address a type of data that has been widely viewed as providing especially strong support for van Riemsdijk’s approach to TFRs, but which, upon closer consideration, turns out to create a hitherto unnoticed problem for it. In a number of studies, van Riemsdijk discusses the implications of a kind of example that constitutes the Dutch counterpart of (15b), and is illustrated in (18a); for completeness, we provide an analogous German example in (18b). Much as with respect to the English data in (13), some speakers of Dutch and German find data like (18) easier to accept if the TFR is preceded by a conjoined adjective, but in order to avoid unintended complexities, we confine our discussion to data as in (18).

(18) a. Bill ontdekte een wat ik zou noemen eenvoudig-e oplossing.
   "Bill discovered a what I would call simple solution."

   b. Bill entdeckte eine was ich nennen würde einfach-e Lösung.
   "Bill discovered a what I would call simple solution."

These data are syntactically different in the following way: While the pivot in (18b), and correspondingly the TFR, are adjectival, the pivots (and the TFRs) in (19a-b) are NPs and DPs respectively. Furthermore, the pivot is differently ordered with respect to the subordinate verb, a matter of importance, as will be seen below. In sum, our principal reason for bringing up the data

---

5 In Grosu (2016), this example was starred. We are grateful to Klaus von Heusinger for pointing out to Alex Grosu that it is acceptable, at least, with a parenthetical intonation.
in (19) is to make clear that such data, although synonymous with (18b), are not directly relevant to the point that will now be discussed.

We now turn to the presentation of a number of grammatical facts which, taken together, provide strong *prima facie* support for the view that the pivot in data like (18) needs to (also) be syntactically realized in the matrix.

(i) First, Dutch and German, as well as a number of other languages that include English, French and Romanian, are subject to the so called Head Final Filter (HFF; Williams 1982), which requires that a pre-nominal AP modifier end with its A head. This condition is straightforwardly fulfilled in (18) if the boldfaced adjective is a matrix sister of the following noun, but appears to be violated if the adjective is assumed to be relative-internal. This problem is not specific to Dutch and German data like (18), it also arises with respect to English, French and Romanian data like those in (15)-(17). The conjunction of the next three properties, however, is specific to Dutch and German.

(ii) A second relevant fact is that in Dutch and German, unlike in English, the head of an AP which modifies a noun must exhibit morphological agreement with the noun (with few exceptions in Dutch, which need not concern us here). Note that in (18), the pivot agrees with the ensuing noun, and this is straightforwardly accounted for if the pivot is the (quasi-)head of the TFR, but unexpected if it is not.

(iii) A third relevant fact is that in Dutch and German, an AP internal to a subordinate clause may not follow that clause's verb, as illustrated in (20b)-(21b); additional illustration is provided in (22)-(23). Now, in (18), the pivot follows the subordinate verb. This is unproblematic if the pivot is assumed to be in the matrix, but unexpected if it is assumed to be in the relative.

(20) a. Jan vraagt zich af [wiek Marie [tk eenvoudig] noemt].

   Jan asks self off who Marie simple calls

   'Jan wonders who Marie calls simple.'

   b. *Jan vraagt zich af [wiek Marie noemt [tk eenvoudig]].

(21) a. Hans fragt sich, [wen Maria dumm nennt].

   Hans asks himself who Mary stupid calls

   'Hans wonders who Mary calls stupid'

   b. *Hans fragt sich, [wen Maria nennt dumm].

(22) a. Deze oplossing zou ik interessant noemen.

   this solution would I interesting call

   'This solution, I would call interesting'

   b. *Deze oplossing zou ik noemen interessant.


   this solution would I interesting call

   'This solution, I would call interesting'

A fourth relevant fact is that in Dutch and German, adjectives must agree with the modified noun when they are in adnominal position, but must not agree with the subject of a copular sentence or small clause when they serve as predicate of that sentence/small clause. Now, if the boldfaced adjective in (18a-b) is in the relative clause, it is the predicate of a small clause selected by noemen/nennen, and the fact that it bears agreement morphology is unexpected. On the other hand, if it is grafted from the (small clause within the) relative clause unto the matrix, the resulting structure has the properties schematically indicated in (24).

\[
(24) \quad [\text{Matrix Clause} \quad \text{Bill ontdekte een eenvoudig-e oplossing} \quad \text{Bill detected a simple-AGR solution} \\
\quad [\text{cp wat ik zou [sc te eenvoudig] noemen}] \quad \text{what I would simple call}]
\]

The two clauses are joined only by the **adjectival root**, which is shared by the relative and the matrix (in keeping with the schema in (3)). The realization of the pivot in the matrix is adnominal, and can serve as basis for the attachment of an agreement affix, as well as for satisfaction of the HFF: the realization of the pivot in the relative is in predicate position, and thus requires no agreement morphology. Furthermore, the (unpronounced) realization of the pivot in the relative may be assumed to precede the subordinate verb, in keeping with requirement (iii) (see (20)). In this way, all four requirements listed above appear to be satisfied under the grafting approach, but seem to be violated under the alternative approach proposed by Grosu (2016).

Before accepting this conclusion, however, it is necessary to check the following prediction made by the analysis in (24) with respect to (18a-b): In view of the fact that English, French and Romanian allow not only constructions in which the speaker subscribes to the content of the pivot, as in (15b), (16a) and (17a), but also constructions in which the speaker disagrees with the content of the pivot, as in (15c), (16b) and (17b), it is expected that Dutch and German, which allow constructions of the former kind, as was seen in relation to (18), should also allow constructions of the latter kind, as in (25).

\[
(25) \quad \text{a. Hij heeft een wat ik niet bepaald zou noemen interessant-e theorie voorgesteld.} \\
\quad \text{He has a what I not exactly would call interesting theory proposed} \\
\quad \text{‘He proposed a what I would not exactly call interesting theory’} \\
\quad \text{b. Er hat eine was ich nie nennen würde einfach-e Lösung vorgeschlagen.} \\
\quad \text{He has a what I never call would interesting solution proposed} \\
\quad \text{‘He proposed a what I would never call interesting theory’}
\]

We submitted the Dutch data in (18a) and (25a) to the evaluation of eight linguistically sophisticated native consultants, and the German data in (18b) and (25b), to seventeen comparably sophisticated native consultants. The consultants were asked to compare the relative acceptability of (18) vs. (25), and the German consultants were also asked to compare (18b) and (25b) pronounced with a continuous vs. a parenthetical intonation, so that (18b) and (25b) got compared with (26a-b) respectively.
(26) a. Bill entdeckte eine, was ich nennen würde, einfach-e Lösung.
   ‘Bill found a, what I would call, simple solution.’

b. Bill entdeckte eine, was ich nie nennen würde, einfach-e Lösung.
   ‘Bill found a, what I would never call, simple solution.’

The following results were obtained:

[A] With respect to (18) and (25) without a parenthetical intonation, our consultants overwhelmingly rated the latter worse than the former (7 of the Dutch, and 15 of the German consultants). Only one Dutch and two German consultants found both types of data OK.

[B] With respect to (18b) vs. (26a), 13 consultants preferred (26a), one consultant expressed no preference, and 3 consultants preferred (18b).

[C] With respect to (25b) vs. (26b), 10 consultants preferred (26b), one consultant preferred (25b), and 6 rejected both.

What do the results [A]-[C] imply for the two competing theories, given the requirements [i]-[iv] noted earlier?

For the grafting approach, [A] constitutes a problem, because the analysis in (24) predicts that both types of data should be OK, [i]-[iv] being satisfied in both. – [B] is compatible with this approach. The preference for a parenthetical intonation manifested by most German consultants may be attributed to the fact that it makes the construal of the pivot in the matrix more salient. – Concerning [C], we do not see how the approach under consideration can shed light on the preference of most consultants for the parenthetical intonation.

For the theory in Grosu (2016), [A] is accounted for rather straightforwardly: Since the pivot needs to be construed in the relative clause, it is an incontrovertible TFR, and the pivot must thus be also syntactically present in the relative only. This state of affairs violates the word-order requirement [iii], illustrated in the (b) sub-cases of (20)-(23), with resulting degradation. [C] can also be accounted for: the intonational separation of the pivot from the subordinate verb plausibly 'masks', or renders more tolerable, the violation of requirement [iii]. This suggestion receives independent support from the observation that Dutch and German typically disallow DPs after a subordinate verb, but nonetheless allow this state of affairs when the post-verbal DP is sufficient long and heavy to induce intonational separation from the verb (see Grosu 2003, examples (11)-(13))

6 Klaus von Heusinger pointed out to us that (25b) can be rendered acceptable by replacing nennen with sagen in it, as in (i).

(i) Er hat eine was ich nie sagen würde einfach-e Lösung vorgeschlagen.
This example is differently structured: while nennen selects the small clause [t einfach], sagen selects a nominal direct object that seems to have quotational status, "einfache Lösung”. If so, (i) does not cast doubt on the unacceptability of (25b).

7 For convenience, we reproduce Grosu's example (13) as (i) below:

(i) Der Hans will der Maria zurückgeben – dieses Buch, diese Platte und diese Kleider.
   the Hans wants the.Dat Maria return this book this record and these clothes
   ‘Hans wants to return to Maria this book, this record and these clothes.’
the matrix, as suggested in the preceding paragraph, the theory need not say anything about the constructions at issue, since as far as it is concerned, these constructions are not TFRs.

It remains to show how Grosu's theory can deal with the judgments of those consultants who found (25a-b) and/or (26b) acceptable, given the requirements (i)-(iv). A first thing to note is that we have so far discussed the requirements (i)-(iv) as if they all had the same status. But it is well known from other domains of study that the relative strength of conflicting requirements may vary from language to language, and sometimes from speaker to speaker, resulting in situations where a construction is acceptable even if it violates a requirement felt to be 'weaker', so long as it satisfies a 'stronger' one (this state of affairs is in fact the primary factor responsible for Optimality Theories). There is independent evidence that the agreement requirement is sometimes stronger than the HFF. For example, the English example in (27) is felt to be at best marginal, due to a violation of the HFF (note that the AP ends with an adjective that is not its head). In contrast, the Dutch and German constructions in (28) are felt to be basically OK, due to the fact that the non-head AP-final adjective bears agreement morphology. The strength of the agreement requirement is also reflected in data like (29), where agreement in AP-final position improves acceptability, even when realized on an AP-final non-adjective.8

(27) ??Mary is looking for an as fast as possible car.

(28) a. Een zo snel als moegelijk-e auto …
   'As fast a car as possible.'
   b. Ich bitte um die [so schnell wie Ihnen möglich-e] Beantwortung meines Briefes.
   'I ask for the as fast as you-Dat possible reply my letter-Gen'
   c. Een zo snel als moegelijk-e auto …

We conjecture that for speakers who find data like (25a-b) OK or marginally possible, satisfaction of (ii), i.e., the agreement requirement, is strong enough to 'mask' the non-satisfaction of (i), (iii) and (iv), i.e., of the HFF, of the word-order requirement, and of the ban on agreement morphology on a predicative adjective. As for the mechanism due to which the pivot of an adjectival TFR receives phi-features from the noun modified by the adjectival TFR, Grosu (2016, p. 1261) suggests that the transparency channel used for accounting for matching effects in category and syntactic number (see section 2) can be exploited in the converse direction to ensure agreement, as alluded to in footnote 4. This mechanism can, of course, be used to analyze adjectival TFRs in any language that requires adnominal adjectives to agree with modified nouns, whether predicative adjectives get inflected or not.

In order to further check our assumption that what makes (25) unacceptable for a significant number of speakers is the violation of a word order requirement, we proceeded to test comparable data in a language in which the agreement requirements are just as in Dutch or

8 For the original discussion of these data in Dutch see van Riemsdijk (1998).
German, but the basic word order in subordinate clauses is VO, not OV. Such a language is Yiddish, and we elicited the data in (30) and (31) from two distinct consultants, who are also native speakers of English and French respectively. Their evaluation of these data was that they have precisely the acceptability of the corresponding data in English and French. This shows that when they are not 'masked' by word order considerations, the relevant data work as predicted by Grosu's transparency-channel mechanism.

(30) a. Bil hot forgeshlag a (nay-em un) vos ikn volt gerufen poshet-n plan.
    Bill has proposed a new-Agr and what I would call simple-Agr plan
b. Bil hot forgeshlag a (nay-em un) vos keyner volt nisht gerufen poshet-n plan.
    Bill has proposed a new-Agr and what nobody would Neg call simple-Agr plan

(31) a. koydemkol vil ikh zogen az
    first-of-all, want I say that
    in meshekh fun di fargangene khadoshim, hobn mir dergreykht
    in the course of the past months, have we achieved
    alts interessanter rezultatn: nor nekhtn, iz bil sofklsof geven mesugl
    increasingly interesting results: only yesterday, was Bill at last in a position
    fortsushlogn a [vos me volt gekent afile rufn posheter-e] farentferung
to propose a what one would be-able even call simpler solution
'I first want to say that in the course of last few months, we have achieved increasingly
interesting results: Yesterday only, Bill was at last in a position to propose a what one
could even call simpler solution.'
b. koydemkol vil ikh zogen az
    first-of-all, want I say that
    in meshekh fun di fargangene khadoshim, zeynen mir tsum tsar nisht
    in the course of the past months, were we unfortunately not
given mesugl tsu dergreykhen
    capable to reach
    keyn vertikl rezultatn: nor nekhtn, hot bil nokh a mol forgeshlogn
    no valuable results: only yesterday, has Bill once more proposed
    a [(farkrimte un) vos keyner bam fulen zinen volt nisht gekent rufn kreativ-e]
a [(screwy and) what no one in his right mind would not be-able call creative]
farentferung
    solution
'First of all, I want to say that in the course of the last months, we have
unfortunately been unable to reach any interesting results: Yesterday only,
Bill proposed once more a (screwy and) what no one in his right mind could
call creative solution.'
4 Summary of Results

In Grosu (2016, section 7), a number of conclusions were reached concerning the relative (de)merits of the two compared theories of TFRs. What has been shown in this paper points to the following additional conclusion: the grafting theory of van Riemsdijk faces problems in connection with the analysis of pre-nominal adjectival TFRs that the theory in Grosu (2016) can avoid.

Acknowledgments

The authors wish to express their gratitude to the audience at IATL 34, for comments made during and after the discussion section, and most especially to Klaus von Heusinger, without whose penetrating remark, noted in footnote 5, this paper would never have been written. We also wish to thank Daniel Birnbaum and Jean Lowenstamm for kindly constructing the data in (30) and (31) respectively. Finally, thanks are due to the numerous Dutch and German consultants who kindly and promptly evaluated the data we submitted to them, as well as to Radek Simik, who enabled us to contact some of them.

References


Riemsdijk, Henk van. 1998a. Trees and scions, science and trees, Chomsky 70th Birthday Celebration Fest Webpage.


THE DIFFERENCE BETWEEN BLENDS AND CLIPPED COMPOUNDS

CAMIEL HAMANS
University of Amsterdam/Adam Mickiewicz University Poznań
CIPL, Comité International Permanent des Linguistes

Abstract
This paper wants to refute the traditional claim that blends or portmanteau words are unsystematic.

First, it will be shown that one must distinguish between two types of concatenations of portions of two words. One the hand formations that combine the first portions of the two source words and on the other words in which the first part of the first source word is combined with the final part of the second. These last group are real blends. The first one is better called clipped compounds, complex clippings or stub compounds.

Both groups show a righthand head. Clipped compounds appear to be a subcategory of compounds and follow the Compound Stress Rule. In blends the right part of the final form is also the head. However, blends copy the prosodic and syllabic structure of the second source word. Whereas compounds consist of at least two prosodic or phonological words, blends consist of only one. This leads to the conclusion that blends can best be described as an intermediate category between compounds and simplex words.

Most of the examples described in this paper come from English, however, some German and Dutch examples are also discussed. Blending operates in a similar way in these languages.

1 Introduction

According to Bauer, Lieber and Plag (2013, 458) there are two types of blends. Both types combine portions of two source words. The first type “involves the loss of medial segmental material” (Bauer, Lieber and Plag 2013, 458) or to put it differently, combines the first part of the first source word, the left source word, and the last part of the second, the right one:
In the second type final segmental material of both source words is truncated:

- modulator + demodulator → modem
- picture/pix + element → pixel
- situation + comedy → sitcom
- frozen + yoghurt → froyo

“The two patterns can be formalized as in (3), where AB stands for the left base (with its two parts A and B) and CD stands for the right base, with its two parts C and D” (Bauer, Lieber and Plag 2013, 458).

(3a) AB + CD → AD
(3b) AB + CD → AC

The formalization of (3a), which says that the first part of the first source word must be combined with the final part of the second source word, leads to blends tout court. The pattern of (3b), where the first portions of the source words are combined, results in what are called ‘clipped compounds’ or ‘clipping compounds’ (Bauer, Lieber and Plag 2013, 458) or ‘stub compounds’ (Spencer, 1998 and Hamans 2018a). Hamans (2018a) prefers the term stub compounds, since stubs are mainly bound elements whereas clippings may appear as free forms (see Hamans 2018b). Because of this reason the term stub compound is used here.

In the remaining of this paper it will be shown that the difference between AD and AC concatenations corresponds with other formal features and that both types of concatenations are systematic.

2 Stub Compounds

Gries (2006) analyses different kinds of ‘subtractive word formation’ from a cognitive linguistic perspective. He convincingly demonstrates that there is a difference between blending and stub compounding. Based on a very sophisticated statistical analysis he is able to show that stub compounding is much less productive than blending. Furthermore, his figures demonstrate that stub compounds systematically preserve less material of their respective source words than blends and finally that the source words of blends are phonologically or orthographically more similar to each other than the source words of stub compounds. In addition, blends may even show an overlap as in slanguage from slang and language, glasphalt from glass and asphalt, motel from motorist and hotel or froogle from frugal and google, whereas stub compounds hardly do. However, the differences Gries found are so heterogeneous that they offer little scope for further systematic research.

Another difference, already noticed by Bauer (1983, 233), is more promising. It is the difference in stress pattern. Stub compounds follow the Compound Stress Rule, whereas blends

Stub compounds get stress on the left part of the resulting AC form, even when stress is on the second lexeme in the full, non-truncated, sequence of source words, as will be shown in the adjective + noun phrases of (4a). Stress assignment in stub compounds also leads to stress on the leftmost part of the final AC form when stress falls on for instance the second or third syllable of the first source word, as shown in (4b).

(4a) mísper < missing pérsón
    mídécult < middle culúté
    fróyo < frozen yóghurt

(4b) sitcom < situatión cómedy
    biopic < biográphý/biográphical pícure
    cýborg < cybernétic órgánism

The examples in (4a) show that these AC formations also behave like compounds in another respect: they appear to follow the Righthand Head Rule (Williams 1981) and thus is the final form a noun, whereas the combination of source words consists of a sequence of an adjective plus a noun. In this respect stub compound can be compared to normal adjective + noun compounds such as greenhouse, bluebird and redhead.

The data presented above show that AC concatenations really behave as compounds, albeit that the constituent parts are not free forms but clippings. The question which now remains is how to clip or truncate the source words. Hamans (2012 and 2018a, b) discusses clipping extensively. The most frequent form of clipping is back clipping, which results in (C)(C)V(C)(C) forms such as:

(5a) tram < tramway
    pic < picture
    ad < advertisement

However, a more recent trochaic pattern also exists:

(5b) psycho < psychopath
    dipso < dipsomaniac
    info < information

Almost all clippings which appear in the examples (4a) and (4b) follow the main (C)(C)V(C)(C) pattern, only bio in biopic is an example of the recent trochaic pattern. This is not the place to discuss clipping extensively, it suffices to show that the clippings of the AC formations discussed here follow standard clipping patterns. The only difference between standard clipping and the clipping process operating in (4a & b) is that standard clipping may result in free forms, whereas the clipping process that operates here only seldom leads to a possible free form. Bio and pic are the only possible free form so far.
3 Head of Blends

According to Gries (2012) blends, the AD type concatenations, have a head, which is usually the right part, which is the remnant of the second source word. Gries uses his semantic criteria and statistical data to reach this conclusion. However, the AD type also exhibits a formal head, as can be seen in the examples (6) and (7).

(6a) Dutch

- *het potel* (n) ‘hotel for Polish workers’ < *de Polen* + *het* hotel
- *het preferendum* (n) ‘referendum with more options’ < *de preferentie* + *het* referendum
- *de scheit* (common gender) ‘mixed breed of sheep and goat’ < *het* schaap + *de* geit

(6b) German

- *der* Kurlaub (m) ‘cure vacation’ < *die* Kur + *der* Urlaub
- *das* Sportel (n) ‘sport hotel’ < *die* Sport + *das* Hotel
- *die* Datei (sg.) ‘file’ < *die* Daten (pl.) + *die* Kartei (sg.)

(7a) English

- *simulcast* (N) < *simultaneous* (Adj) + *broadcast* (N)
- *malware* (N) < *malicious* (Adj) + *software* (N)
- *barkitecture* (N) ‘design of doghouses’ < *bark* (V) + *architecture* (N)

(7b) German

- *Naktivist* (N) ‘naked activist’ < *nackt* (Adj) + *Aktivist* (N)
- *herrklären* (V) ‘mansplain’ < *Herr* (N) + *erklären* (V)
- *Teuro* (N) ‘nickname for the expensive Euro’ < *teuer* (Adj) + *Euro* (N)

(7c) Dutch

- *vagetariër* (N) ‘vague vegetarian’ < *vaag* (Adj) + *vegetariër* (N)
- *krommunicatie* (N) ‘crooked communication’ < *krom* (Adj) + *kommunicatie* (N)
- *alterneut* (N) ‘an unqualified healer’ < *alternatief* (Adj) + *therapeut* (N)

The examples in (6) and (7) show that the righthand part determines the gender, the number and the part of speech of the blend. The examples (6a) are simple and clear. In Dutch there are two genders, neuter and common gender. The gender of the second source word determines the gender of the blend. The German data of (6b) are somewhat more complicated. German has three genders, male, female and neuter. Again, it is the gender of the second source word which is decisive for the gender of the blend. The same applies to number as the example *Datei* (6b) shows.

In (7b & c) the second source word determines the resulting part of speech. When the second source word is a noun, then the resulting blend also is a noun, whatever the part of speech of the

---

1 Examples from Pajerová (2018)
2 Even the orthography of the German examples shows which part is the head. In *Naktivist* and *Teuro* the original adjectival parts *n(akt)* and *teu(er)* receive substantive Großschreibung ‘capitalization’ since the resulting blend is a substantive due to the second source word. In *herrklären* it is just the other way around: the substantive *Herr* has to give up its capital letter since the blend is a verb due to the verbal character of the righthand part, the head.
first source word is. However, when the second source word is a verb, as in *herrklären*, then the final blend is also a verb.

These examples show that blends exhibit a formal head, just as compounds. However, this does not imply that all blends must also have a semantic head. Just as dvanda compounds do not show a semantic head, see for instance *singer-songwriter*, *bittersweet* and *spacetime*, where the meaning is the sum of the meanings of the two constituent words, ‘copulative’ or dvanda blends such as *smog*, from *smoke* and *fog*, or *brunch* from *breakfast* and *lunch* or *Oxbridge* from *Oxford* and *Cambridge* do not exhibit a semantic head. However, all blends have a formal head, and, in this respect, blends behave as compounds.

4 Prosodic Aspects

This section shows how blends copy the prosodic and syllabic structure of the second source word. First stress assignment will be discussed. The second part of this section is devoted to syllable structure and actually discusses which parts of which source word can be combined.

4.1 The Stress Pattern of Blends

Beard (1998:57) was the first to observe that the prosodic structure of blends must be identical with that of the model, being the second source word. To put it differently: blends tend to copy the stress pattern of the head (see also Piñeros 2000 & 2002, Bat-el 2006, Bat-el and Cohen 2012, Trommer and Zimmerman 2012). The examples in (8) demonstrate the stress pattern of blends.

(8) boatél < boat + hôtél
frappucíno < frappé + cappuccíno
flustátéd < flústered + frustráted
advertóríal < advértisement+ éditóríal
fertigátión < fértílizer + irrigátión
préstínant < prestúgious + dominant

In all these examples it is the stress pattern of the second source word that determines the stress placement on the resulting blend. The last example, *préstínant*, is most convincing. Even when the segmental material of the second source word is not preserved, the suprasegmental prosodic feature stress of this source word retains its strength. The resulting blend bears stress at exactly the same place as the second source word, on the first syllable, notwithstanding the fact that this syllable was unstressed in the first source word.

Blends appear to consist of one prosodic word, although they are formed as a concatenation of parts of two separate words. The prosodic shape of a blend is the same as that of the second source word. Blends simply copy the primary word stress of the second source word.

---

3 The notion prosodic or phonological word stands in opposition to the notion grammatical word (Booij 1999,47). “Prosodic words are typically characterized as being the domain of word stress, phonotactics and segmental word-level rules” (Peperkamp 1999, 15). The size of the prosodic or phonological word does not have to correspond with the morphological word. For instance, compounds in English, German and Dutch consist of two prosodic words. (See about prosodic words also Peperkamp 1997 and Hildebrandt 2015)
4.1.1 Monosyllabic Source Words

Bat-El and Cohen (2012) discuss the relation between blending and stress assignment in English in detail. They claim that two factors play a role in determining the position of stress in blends. The first one is position, the second size. Here it will be argued that position will do for almost all data.

Bat-El and Cohen (2012) agree that the main pattern of stress placement in blends is a copying process of the prosodic structure of the second source word as in (8), which means that the stressed syllable of the blend is identical to that of the second source word. Stress is position-based in these cases. However, there are exception they show, such as blends with a monosyllabic source word. For these blends size should determine stress assignment.

(9) blógive < blog + árchiple
    tankíni < tank + bikíni
    momprenéur < mom + entreprenéur

(10) lúmist < lúminous + mist
     citrisun < citric + sun
     éscalift < éscalator + lift

The blends in (9) simply follow the stress pattern of the second source word and and should therefore not be considered as exceptional or as counterexamples. The data presented in (10) does not indeed follow the stress pattern of the second source word, which, incidentally, is completely predictable. The second source words are monosyllabic words and thus have no lexical stress (Bat-El and Cohen (2012: 207) or metric pattern or rhythmic contour of their own. Consequently, the resulting blend has to copy the only available stress pattern or rhythmic contour, which is the pattern of the first, left, source word. There is no reason to take size as a determining factor in these cases.

However, there are a very few real counterexamples

4.1.2 Counterexamples

Usually the first source word of a blend contains fewer syllables, and is therefore shorter, than the second one (Kelly 1998). However, this is not a condition as brunch, from breakfast + lunch, demonstrates. When the size of the second source word is smaller than that of the first source word (sw2<sw1), exceptions to standard blend stress assignment may occur, as Bat-El and Cohen (2012) show. In (11) the resulting form adopts the stress pattern of the first, left, source word, whereas in (12) the stress pattern of the second source word is copied.

(11) húrricoon < húrricane + ballóon
    hándkerchoo < hándkerchief + kercchóo
    quálatex < quálity + látex

---

4 Size is considered to be the main factor for stress assignment in blends by Cannon (1986).
5 Examples (9) and (10) taken from Bat-El and Cohen (2012)
6 Examples (11) and (12) taken from Bat-El and Cohen (2012)
Bat-El and Cohen (2012:202) conclude on the basis of these data that there is a certain ‘inter-word variation where different words follow minimally different rankings’ of constraints. However, they need these different constraints because of the behaviour of blends with a monosyllabic source word, discussed above. As shown in 4.1.1, only a small group of blends resulting from monosyllabic second source words (cf. 10) does not follow standard blend stress assignment. Here the outcome is the default option. Therefore, monosyllabic source words do not require a size constraint.

The examples in (12) simply follow the standard blend stress assignment. So, only the data in (11) might be considered to contradict the normal stress copying pattern. However, some of the data in (11) are not very convincing: *handkerchoo* and *qualatex* look more an AC concatenation than an AD. They are better described as a sort of stub compounds. Real counterexamples are blends such as *hurricoon* and *ballute*:

(13) húrricoon  < húrricane + ballóon
    ballúte   < ballóon + párachute

It should be noted that in *ballute* the first source word is smaller in size than in the second.

For the right stress placement Bat-El and Cohen (2012) suggest a few faithfulness constraints that preserve the phonological properties of the base words at the segmental level as well as at the level of metrical structure. For the position-based view of stress assignment they suggest two constraints that state that the stressed syllable in the blend corresponds to the stressed syllable in the respective source word. In order to put the stress on the correct syllable of the right word, both constrains are ranked in the following way.

(14) FAITHHEADWR >> FAITHHEADWL

Since blend stress normally corresponds to that of the right constituent of the blend, the candidate that does not violate stress assignment required by FAITHHEADWR wins. For the size criteria of stress placement Bat-El and Cohen (2012: 199) suggest another constraint: FAITHMETRICALSTRUCTURE (FAITHMS), which states that ‘[t]he metrical structure (number of syllables and stress pattern) of the blend is identical to that of both base words’. Different rankings of FAITHMS in relation to FAITHHEADWR and FAITHHEADWL account for the difference in stress assignment.

However, the constraints Bat-El and Cohen propose can easily be simplified. As demonstrated, only blends with a monosyllabic second source word violate systematically FAITHMS, provided that this constraint is split into FAITHMSWL and FAITHMSWR. In order to produce blends with correct placement on the right source word, the ranking of these two constraints must be:

(15) FAITHMSWR >> FAITHMSWL
Since the second source words in the examples of (10) do not have any lexical stress or rhythmic contour, faithfulness to the metrical structure of the second source word (WR) is vacuous. Consequently, the resulting blend remains faithful to the first source word (WL). The only real counterexamples are blends such *ballute* and *húrricoon*. These blends show the stress pattern and the syllabic skeleton of the first source word.

These examples do not belong to a single category. The source words of *ballute* follow the normal pattern: sw2>sw1. In *húrricoon* it is just the other way around, which is exceptional. One may try to explain the exceptional behaviour of these examples by pointing to the exceptionally large portion that is deleted in *parachute*, whereby even the place of stress appears to be erased, just as in *balloon*. Because of so many and serious violations the contour of the first source word may get priority. However, this explanation sounds rather ad hoc, when one realises that in examples such as (16) the phonemic content of the stressed syllable is deleted without any consequences for the stress pattern.

\[(16) \text{blógive} \quad \text{< blog} \quad \text{+ árchive} \]
\[\text{préstinant} \quad \text{< prestigious} \quad \text{+ dómínant} \]
\[\text{plúmcot} \quad \text{< plum} \quad \text{+ ápricot} \]

Therefore, it seems better to accept that there is a very small group of exceptions, of which most show a difference in source-word length which is exceptionally sw2<sw1. In this group an opposite ranking applies, which means that there are two rankings available in English, of which (15) is the preferred one. However, (17) also exists, which implies that there is a ‘crucial non-ranking’ between FAITHMSWR and FAITHMSWL.

\[(17) \text{FAITHMSWL} >> \text{FAITHMSWR} \]

The fact that there are two possible rankings, of which one is the preferred one, is not exceptional (cf. Hamans 2012 on two possible rankings for Dutch clippings of which one is the preferred ranking).

### 4.2 The Syllabic Structure of Blends

Stress assignment is not the only aspect which blends copy from their second source words. Usually, the syllabic structure of blends is also a copy of the syllabic structure of the second source word as the examples in (18) show (see also Arndt-Lappe and Plag 2012).

\[(18a) \text{breakfast} \quad \text{+ lunch} \quad \rightarrow \text{brunch} \quad \text{onset} \]
\[\text{smoke} \quad \text{+ fog} \quad \rightarrow \text{smog} \quad \text{onset} \]
\[\text{boat} \quad \text{+ hotel} \quad \rightarrow \text{boatel} \quad \text{onset}^7 \]

---

7 An alternative segmentation for this example may be, especially when one wants to give priority to the orthographic form:
The examples presented here demonstrate that it is the syllabic structure of the second source word which determines the syllabic structure of the blend. If only an onset has been truncated from the second source word, then only an onset can be inserted. However, it should be noted that empty onsets can be truncated and refilled as in *glasphalt*, where the first syllable of the second source word *as* does not include an onset. A similar insertion applies to *donkephant* from *donkey* and *elephant*. The syllable *el* without onset is replaced by a syllable with an onset *donk*.

In addition, instead of a one place coda the first syllable of the blend exhibits a two-place coda, which does not make any difference in terms of syllabic structure. It also appears possible to fill a one place onset with a cluster consisting of two or three consonants as in shown in the series *glitterati* from *glitter* + *literati*, *clitterati* from *clitoris* and *literati* and *splitterati* from *split* and *literati*. This is not in any way contrary to the possibilities described here, since English (and Dutch) onsets can consist of more than one consonant.

If truncation leads to the deletion of a syllable, the resulting empty space must be filled with a syllable as shown in (18c). However, it is also possible to insert more syllabic material as in the following examples of Dutch:

(18e) *anachronisme* + *acroniem* → anacroniem

‘acronym derived from an outdated phrase (e.g radar)’

---

*(18b) Greek/Greece + exit* → Grexit

*Span*ish  + *English*  → Spanglish

*gin*gantic + *enormous* → ginormous

(18c) *stegan*ation + *inflation* → stagflation  σ (= syllable)

*Ox*ford + *Cambridge* → Oxbridge  σ

*gue*s + *estimate* → guesstimate  σ^8

(18d) *ad*vertisement + *edit*orial → advertorial  σσ^9

*edu*cation + *entertain*ment → edutainment  σσ

*stalker* + *paparazzi* → stalkerazzi  σσ^10

*boat* + *hotel* → boatel

In this case a whole syllable consisting of an onset and a nucleus have been truncated from the second source word. Subsequently, the onset and the nucleus of the first source word are inserted.

Another alternative might be:

*boat* + *hotel* → boatel.

In this case a whole open syllable plus the onset of the next syllable must have been truncated. Consequently, a CVC syllable must have been inserted, of which the last consonant has been resyllabified in order to become the onset of the final syllable. However, examples such as Dutch *potel* ‘hotel for Polish immigrant workers’ from Polen + *hotel* or *stotel* ‘student hotel’, without final -t at the end of the first source word, make this last segmentation unlikely.

^8 An alternative segmentation may be:

*guess* + *estimate*.

In this case only the empty onset of the first syllable of the second source word has been truncated and replaced by the onset of the first source word.

^9 An alternative segmentation could be:

*advertisement* + *editorial* → advertorial.

However, a form such as *prefatorial*, from *preface* without a -t makes this segmentation unlikely.

^10 An alternative segmentation may be:

*stalker* + *paparazzi* → stalkerazzi.

Also, in this case the deletion and insertion will affect two possible syllables.
democratie + dictatuur → democratuur
‘dictatorship by a democratic chosen leader’

In *acronym* the first syllable *a* is truncated. Subsequently two syllables, *a* + *na*, are inserted. The second example *democratuur* is even more complicated. In *dictatuur* the first syllable *dic* is truncated plus the onset of the following syllable *t*. Subsequently two syllables *de* + *mo* plus a following onset *kr*, spelled out as *cr*, have been inserted. The extra syllables *a* and *de* in *anacroniem* and *democratuur* can be described as unparsed. In fact, such an extra, unparsed, syllable does not affect the overall picture that blends copy the syllabic structure of the second source words, of which the remnant becomes the formal head of the blend. Also, in this respect the blend remains one phonological word.

A problem that will only be touched on here is where to cut off the second source word. So far, it is clear that what has been cut off should be supplemented from the initial segment of the first source word. However, the decision whether an onset, an onset plus nucleus or a syllable or even more should be truncated seems arbitrary, although it is evident that one cannot delete a whole syllable from a monosyllabic source word. Gries (2006) shows that the recognizability or recoverability of the original source word or of similar lexical competitors – forms with a similar form and meaning – plays an eminent role in the selection of the cut-off point. Hamans (2010) points to Zabrocki’s theory of diacrisis for an explanation. Zabrocki (1962 and 1969) expands the notion of minimal pair, he compares segments that differ in more respects than one phoneme or feature. He calls the corresponding parts of lexemes confusive segments or confusiva. For instance, the lexemes *crack* and *pack* which are not a minimal pair since the first one starts with a consonant cluster, whereas the initial segment of *pack* is only one consonant, share a confusivum *ack*. A successful blend must contain confusiva with both source words that are large enough to trigger recognition of the original form in the mind of the listener.

Since the speaker or the word coiner starts with truncation of the second source word, he is obliged to keep as much material of the second source word so that it can easily traced back to the full original. For a successful blend the confusivum should be large enough to be easily traceable. Subsequently the speaker must fill the truncated syllabic position(s) with corresponding material of the first source word. The segmental material taken from the first source word also forms a confusivum with the original source word and with similar lexical competitors. When this confusivum is not large enough to make it easily traceable the onset may be filled with more consonants or even an unparsed initial syllable may be added. Gries’ (2006) metrical and statistical procedures may be useful to determine when a remaining part can easily be traced back. However, this is a matter for further psycholinguistic research just as the role of the overlap in blends such as *slanguage* from *slang* and *language*.

---

11 Alternative segmentations may be proposed, such as *democratie + dictatuur* or *democratie + dictatuur*. However, this does not make any difference for the argument.

12 One can also describe *anacroniem* as a blend where no truncation of a part of the second source word has taken place. In this case the empty onset of the first syllable of *acronym* is filled by the initial segment *an* from *anachronism*. The result is the same: an extra syllable is added to the blend. Truncation of a segment of the second source word is not mandatory as the examples *slanguage*, *guesstimate* and *glasphalt* discussed before show.
5 Blends as an Intermediate Category

In section 3 it is demonstrated that blends and compounds share the characteristic of the righthand part as their formal head. The head determines the grammatical properties of the blend. Insofar the concatenation of parts of two source words that results in a blend behaves as if it was a compound. In this respect there is no difference with stub compounds. However, phonologically blends cannot be described as a sort of compounds. The constituents of a compound form each a prosodic or phonological word (Booij 1995, 49). So, bluebird and greenhouse consist of two phonological words each just as sitcom and midcult. However, the constituents of a blend together form one phonological word, as shown in section 4. This phonological word is normally a copy of that of the second source words, the base of the righthand part.

Most of the second source words which pop up in blends are underived, monomorphemic thus simplex words. Consequently, most blends can also be described as simplexes from a phonological point of view. Since blends combine characteristics of compounds and of simplex words at the same time, they should be described as an intermediary category. Even blends which have a derived word as second source word do not contradict this observation. All the complex words that can act as second source word and that are presented here contain a vowel initial suffix. Such words form a single prosodic word, as Raffelsiefen (1999) demonstrated.

(19a) -ish
Spanglish < Spanish + English
(19b) -ial
advertorial < advertisement + editorial
(19c) -ity
flexicurity < flexible + security
(19d) -er
compander < compressor + expander
(19e) -ation
fertigation < fertilize + irrigation

The resulting blends form each one phonological word, just as the second source word of which they are a phonological copy. In terms of word formation blends appear to behave as compounds, however from a phonological perspective they consist of only prosodic word.

6 Conclusion

- There is an essential formal difference between AD-blends and AC-clipped compounds, stub compounds. Stub compounds are compounds of two clipped lexemes. Some of these clipped lexemes are already free morphemes.
- Being compounds stub compounds have a righthand head and exhibit the compound stress rule.
- Blends are also concatenations of parts of two source words. However, blends are composed of the initial or left-hand part of the first source word and the final or righthand part of the second source word. These parts are usually not free morphemes.
- Blends also exhibit a formal righthand head, which suggests that blending is a form of compounding.
- However, the compound stress rule does not apply to blends.
- Blends form a single phonological word, which is a copy of the prosodic and syllabic properties of the second source word.
- Therefore, blends can best be described as a borderline or intermediary case between compounds and simplex words, especially in the case of blends derived from a monomorphemic second source word.

References


Hamans, Camiel (2010). “The productivity of blending: linguistic or cognitive? Or how to deal
Blends and Clipped Compounds


TWO TYPES OF OBJECT INCORPORATION IN UZBEK

ZARINA LEVY-FORSYTHE
OLGA KAGAN
Ben-Gurion University of the Negev

1 Introduction

This paper investigates object incorporation in Uzbek (Southeastern Turkic/Karlu), a phenomenon whereby a nominal in the object position integrates into the verb, yielding a closely associated verbal unit much different from a parallel non-incorporated direct object construction. The question of whether Turkic languages employ incorporation and if they do, whether it is true incorporation (TI) or pseudo-incorporation (PI) is subject to an on-going debate. The current study contributes to the discussion by (a) analyzing Uzbek data, under-investigated within the generative framework, and (b) arguing that the same language may exhibit both PI and TI.

The focus of the present work is on Uzbek ‘bare nominals’, i.e. nominals which do not contain determiners, overt quantifiers and inflectional morphology, such as number- and case-marking. Specifically, we argue that, most typically, bare nominals in the object position are pseudo-incorporated (1a) and as complements of light verbs (1b), they may be treated as truly incorporated.

(1) a. Anvar rasm chizdi
    Anvar picture draw.PST.3SG
    ‘Anvar drew (a) picture(s).’
   
b. Anvar rasm soldi
    Anvar picture put.PST.3SG
    ‘Anvar drew (a) picture(s).’ (= ‘Anvar drew.’)
In discussing various properties of incorporated nominals, we compare them to non-incorporated full-fledged direct objects, i.e. nominals carrying various morphological elements, such as determiners, quantifiers, number and case suffixes. For instance, (2a) contains an example of a singular definite object, marked with the accusative case suffix -ni. Direct objects can also contain the indefinite determiner bir/bitta (‘one, a’), which gives rise to an existential reading of the nominal (2b). Such objects may occur either marked or unmarked. As a rule, the presence of the accusative indicates specificity, while its absence signals the nominal’s non-specificity. Lastly, (2c) illustrates a plural object, carrying the plural suffix –lar. While unmarked plural objects are interpreted as indefinite and non-specific, case-marked plurals receive definite interpretation.

(2) a. Anvar rasm-ni chizdi
   Anvar picture-ACC draw.PST.3SG
   ‘Anvar drew the picture.’
b. Anvar bir/bitta rasm(-ni) chizdi
   Anvar one picture-ACC draw.PST.3SG
   ‘Anvar drew a (non-specific) picture.’/ ‘Anvar drew a (specific) picture.’
c. Anvar rasm-lar(-ni) chizdi
   Anvar picture-PL-ACC draw.PST.3SG
   ‘Anvar drew pictures.’/ ‘Anvar drew the pictures.’

In the discussion below, we show that in contrast to the full-fledged regular direct objects as in (2), bare nominals as in (1) exhibit morpho-syntactic and semantic hallmarks of incorporation, such as reduced or minimal nominal structure, obligatory narrow scope, number-neutrality, atypical discourse anaphora and name-worthiness (Dayal, 2003, 2011, 2015; Borik & Gehrke, 2015).

The paper is organized in the following way: section 2 provides theoretical background on object incorporation and presents central notions pertaining to the cross-linguistic phenomenon. In section 3, we explore object incorporation in Uzbek. Specifically, in subsection 3.1 we argue that bare nominals combined with main verbs undergo pseudo-incorporation and in subsection 3.2 we argue that bare nominals combined with light verbs undergo true incorporation. Finally, section 4 concludes the paper.

2 Object Incorporation across Languages

The term ‘incorporating’ (‘einverleibend’) was used earliest by von Humboldt (1836) in relation to polysynthetic languages, such as Nahuatl (Uto-Aztec), in which words can be comprised of multiple lexical roots. The discussion of the phenomenon has originated from and has been central to the study of various Native American languages, with the term having eventually come

---

1 Uzbek, like other Turkic languages, is a differential object marking (DOM) language (see Guntsetseg et al., 2008; von Heusinger et al., 2008).

2 Bir originates from the numeral ‘one’ and bitta is its classified form, containing a Persian suffix –ta. Here we treat them as two forms of the same article, using in the rest of the examples the form bitta, which is more common in Modern Standard Uzbek. However, see von Heusinger & Klein (2013) for the contrasts between the two.
to refer to a specific instance of a coalescence of a nominal in the object position and a verb, known as ‘noun incorporation’ or ‘object incorporation’. In this regard, a seminal debate between Kroebner (1909, 1911) and Sapir (1911) demonstrates that the nature of incorporation has been contested from early on. As Haugen (2008) points out, the essence of this debate relates to a once major theoretical issue: the division of labor between morphology and syntax in word-formation, and whether or not morphology exists as a separate module. Hence, according to Haugen, the debate foreshadows the later dispute between syntactic and lexical approaches to word-formation, and their respective treatment of noun incorporation across languages (see, among many others, Sadock, 1980, 1985; Baker, 1988; Mithun, 1984; di Sciullo & Williams, 1987; Rosen, 1989).

Over the years, the cross-linguistic inquiry into the phenomenon has also gained a semantic dimension, with a number of influential studies, identifying essential semantic attributes of incorporation related to number interpretation, scopal behavior, and anaphoric potential (see among others, Bittner 1994; van Geenhoven, 1996, 1998). Subsequently, object incorporation has come to be viewed as a phenomenon encompassing two equally important components, namely ‘morpho-syntactic’ and ‘semantic’ (Massam, 2001; Dayal, 2003, 2011, 2015; Farkas & de Swart, 2003; Espinal & McNally, 2011; Modarresi & Simonenko, 2007; Modarresi, 2014; Krifka & Modarresi, 2016, Borik & Gehrke, 2015).

Starting with the pivotal work by Massam (2001), the literature distinguishes between two types of object incorporation: true incorporation (aka canonical/classical incorporation) and pseudo-incorporation.

In the broadest possible sense, true incorporation (TI) involves a morpho-phonological and/or a morpho-syntactic fusion of the nominal and the verb, as shown in (3a). Cross-linguistically, TI nominals have a minimal structure, i.e. they are at most N₀s (roots, stems, or head nouns), and share a strong tie with their verbal predicates, appearing strictly verb-adjacent. Frequently TI is characterized by a change in valency of the incorporating verb, the so-called detransitivization, which is especially vivid in ergative-absolutive languages (cf. Mithun, 1984; van Geenhoven, 1998).

(3) Yucatec (Mayan) (based on Mithun, 1984: 857):

a. ɛi'ak- ē'-'n-ah-en
   chop-tree-ANTIPASS-PERF-I (ABS)
   ‘I wood-chopped’ = ‘I chopped wood.’

b. t-in-ɛi'ak-Ø-ah  ē'
   COMP-I-chop-it-PERF tree
   ‘I chop a tree.’

Pseudo-incorporation (PI) comprises a less allied relation between the nominal in the object position and the verb (4a) (Massam, 2001; Dayal, 2003, 2011, 2015; Farkas & de Swart, 2003; Espinal & McNally, 2011; Öztürk, 2005, 2009; Kamali, 2015; Modarresi, 2014). Both TI and PI nominals lack morpho-syntactic elements corresponding to the marking of definiteness/specificity and number. However, unlike in TI, PI nominals allow a variety of

---

3 The term ‘object incorporation’ is used in the literature descriptively to refer to the fact that such incorporation targets the position usually associated with direct objects and to distinguish the phenomenon from ‘subject incorporation’.
phrasal modifiers, e.g. adjectives, participles, relative clauses, and, in some languages, case- and/or number-marking. Additionally, PI nominals exhibit a relative morpho-syntactic independence from their incorporating verbs. They generally lack strict verb-adjacency and do not affect verbal valency.

   a. Ram *macchli pakaR* rahaa hai  (pseudo-incorporation)
      Ram(MASC) fish(FEM) catch PROG-MASC-SG be-PRS
      ‘Ram is catching fish.’
   b. Ram-ne *macchli-ko pakaRaa*   (lack of incorporation)
      Ram-ERG fish(FEM)-ACC catch-MASC-SG-PFV
      ‘Ram caught the fish.’

Importantly, both types of incorporated structures share morpho-syntactic and semantic characteristics, which set them apart from ‘full-fledged/regular direct object constructions’, i.e. constructions, containing a non-incorporated referential DP (determiner phrase) complement of a transitive verb. Broadly, the attributes of object incorporation are:

**I. Morpho-syntactic deficiency:**

a) **Nominal structure:** incorporated nominals lack functional projections pertaining to regular DP/KP objects. While in TI the nominal structure is **minimal**, i.e. bare noun (N$^0$), in PI, it is **reduced**, i.e. phrasal (NP, even NumP). This point is illustrated in the following schema by Dayal (2015):

(5) (from Dayal, 2015: 52):

```
DP ← non-incorporated nominals
  D
NP ← pseudo-incorporated nominals
  #/Modifiers
N ← canonically incorporated nominals
```

b) **Syntactic position:** the distance between the incorporated nominal and the verb is minimized (6). In TI, both elements are completely bare and form a single morpho-syntactic unit. In PI, nominals are closer to verbs than regular direct objects, but have a relatively independent status as syntactic complements of transitive verbs.
II. Semantic deficiency:

a) Obligatory narrow scope: incorporated nominals are scopally inert and cannot receive a wide-scope interpretation.

b) Number-neutrality: morpho-syntactically singular incorporated nominals receive semantically singular and plural interpretations.

c) Atypical anaphora: anaphoric accessibility of incorporated nominals stands in contrast to that of regular non-incorporated objects.

d) Name-worthiness: an incorporated verbal construction is interpreted as a conceptual whole.

Both types of object incorporation have been proposed for Turkic languages. Some of the researchers argue for TI (Mithun, 1984; Knecht, 1986; Kornfilt, 1997, 2003; Aydemir, 2004, all regarding Turkish) and some for PI (Öztürk, 2005, 2009; Kamali, 2015 for Turkish; Baker, 2014 for Sakha) and yet some against both TI and PI (Lyutikova & Pereltsvaig, 2015 for Tatar) in a given language.

The advancement of these competing proposals concerning Turkic languages raises a natural question as to whether object incorporation takes place in Uzbek.

In the next two sections, we demonstrate that both morpho-syntactic and semantic evidence indicates that Uzbek employs incorporation of bare nominals. We argue that the type of incorporation depends on the verb: main verbs are involved in pseudo-incorporation and light verbs, in true incorporation.

3 Object Incorporation in Uzbek

3.1 Pseudo-Incorporation

The investigation of constructions containing a bare nominal and a main verb shows that there is no true morpho-syntactic fusion between the two. Bare nominals (i) have a reduced, but phrasal

---

Lyutikova & Pereltsvaig (2015) argue that Tatar bare nominals are not incorporated, but briefly suggest that bare nominals in light verb constructions may be analyzed as pseudo-incorporated.
structure and (ii) maintain a relative independence, acting as syntactically mobile complements of transitive verbs (similar to full-fledged direct objects).

Furthermore, bare nominals in such constructions exhibit all semantic hallmarks of object incorporation: (i) obligatory narrow scope, (ii) number-neutrality, (ii) atypical discourse anaphora of the nominal and (iv) name-worthiness. We conclude that these properties together serve as evidence against TI and in favor of PI.

Finally, we propose that bare nominals denote properties (of type \(<e, t>\)), that get incorporated into transitive verbs (of type \(<e, <e, t>>\)) by the non-saturating semantic mode of composition, namely predicate restriction (Chung & Ladusaw, 2004; Modarresi, 2014).

3.1.1 Morpho-Syntactic Evidence
As mentioned earlier, morpho-syntactic deficiency is a cross-linguistically acknowledged property of incorporation. In Uzbek, bare nominals are deficient both in their functional structure and in syntactic position.

a) Reduced phrasal nominal structure
Typically, the functional architecture of Uzbek bare nominals is reduced but phrasal. In addition to lacking case-, number-marking and the indefinite article bir/bitta (cf. (2) above), bare nominals cannot be preceded by demonstratives, universal quantifiers (whether collective or distributive) and possessives. In the presence of all of the above, the nominal obligatorily receives accusative marking and is ungrammatical without it (7). The unacceptability of the elements in question indicates a reduced nominal structure.\(^5\)

(7) a. Anvar \textbf{bu} rasm-*(ni) chizdi
   Anvar this picture-ACC draw.PST.3SG
   ‘Anvar drew this picture.’

b. Anvar \textbf{hamma} rasm-*(ni)/ \textbf{har} \textbf{bitta} rasm-*(ni) chizdi
   Anvar every picture-ACC every one picture-ACC draw.PST.3SG
   ‘Anvar drew every picture.’

c. Anvar Ra’no-\textbf{ning} rasm-i-*(ni) chizdi
   Anvar Rano-GEN picture-3SG-ACC draw.PST.3SG
   ‘Anvar drew Rano’s picture.’

At the same time, bare nominal complements of main verbs have a phrasal structure: they allow modification by adjectives (8a) and relative clauses (RC) (8b). For instance, adjectival adverbs like \textit{g’oyibona}, which are ambiguous between adjectival and adverbial meanings, result in the two readings specified in (8a). Thus, \textit{g’oyibona} can modify both the whole verbal construction and – crucially for our purposes - the bare nominal. In turn, RC modification is only possible if the RC serves as a general characteristic of the nominal, without causing it to be

\(^5\) We assume that Uzbek nominal architecture includes the DP projection. However, note that the presence of DP in Turkic languages is debated (see among many others, Lewis, 1967; Kornfilt, 1997; Öztürk, 2005; Bošković & Şener, 2014; Arslan-Kechriotis, 2006 for Turkish; Lyutikova & Pereltsvaig, 2015 for Tatar; von Heusinger & Kornfilt, 2017 for Turkish, Azeri, Uzbek, Kirghiz, Sakha).
interpreted as specific. Thus, in (8b) the RC modifying the bare nominal refers broadly to the quality of Anvar’s picture(s), which is clear from the imperfective aspect.

(8) a. Anvar g’oyibona rasm chizdi
    Anvar secret(ly) picture draw.PST.3SG
    1. ‘Anvar drew (a) picture(s) secretly.’
    2. ‘Anvar drew (a) secret picture(s).’

b. Anvar hammani hayratlan-tir-adi-gan rasm chiz-adi
    Anvar everyone.ACC astonish-CAUS-IMPRF.3SG-PTCPL picture draw-IMPRF.3SG
    ‘Anvar draws (a) picture(s) that astonish(es) everyone.’

The possibility of such modification reflects the potential phrasal structure of such nominals, which is a strong evidence against their TI.

b) Syntactic independence

Cross-linguistically, TI is identified based on the morpho-syntactic fusion of a nominal (N⁰) and a verb (V⁰), frequently resulting in detransitivization of the verb. In Uzbek, however, no such fusion takes place between a bare nominal and a main verb. Although preverbal, such bare nominals do not need to be strictly verb-adjacent and act as syntactically visible complements of transitive verbs. These properties taken together point against TI in the constructions under discussion.

The example below shows that both case-marked and unmarked complements of main verbs may appear as stand-alone answers to a ‘wh’-question. This serves as an indication that such nominals, including bare nominals, are syntactic constituents.

(9) Q: Anvar nima ko’rdi?   A: Bitta kino(-ni)/ kino-ni/ kino
    Anvar what see.PST.3SG one movie-ACC movie-ACC movie
    ‘Q: What did Anvar see? A: A movie/ the movie/ (a) movie(s).’

Identical to that in Turkish (Öztürk, 2009), it is possible to omit the verb under identity (10a), conjoin verbs serving as predicates of the same object nominal (10b) and conjoin bare nominals in the preverbal position (10c).

(10) (adapted from Öztürk’s Turkish example, 2009: 339):

a. Anvar kino emas, spektakl ko’rdi
    Anvar movie not play see.PST.3SG
    ‘Anvar saw (a) play(s), not (a) movie(s).’

b. Anvar rasm chizdi va men-ga berdi
    Anvar picture draw.PST.3SG and I-DAT give.PST.3SG
    ‘Anvar drew (a) picture(s) and gave it (them) to me.’

c. Anvar portret va manzara chizdi
    Anvar portrait and landscape draw.PST.3SG
    ‘Anvar drew (a) portrait(s) and (a) landscape(s).’

As shown in all examples above, Uzbek bare nominals take a preverbal position, which falls into the general Turkic pattern. In neutral sentences, i.e. sentences with default word order and
information structure, bare nominals cannot be separated from the verb (11a), whereas regular objects can (11b):

(11) (adapted from Baker’s Sakha example, 2014: 8):
   a. Men Masha-ga kitob (*Masha-ga) berdim
       I Masha-DAT book (*Masha-DAT) give.PST.1SG
       ‘I gave (a) book(s) to Masha.’
   b. Men Masha-ga kitob-ni (Masha-ga) berdim
       I Masha-DAT book-ACC (Masha-DAT) give.PST.1SG
       ‘I gave the book to Masha.’

However, in Uzbek, unlike in Sakha (Baker, 2014) and Tatar (Lyutikova & Perel’tsvaig, 2015), and similar to Turkish (Öztürk, 2005, 2009; Kamali, 2015) and other PI languages (see among others, Dayal, 2015 for Hindi; Modarresi, 2014 for Persian; Farkas & de Swart, 2003 for Hungarian), adjacency of bare nominals to main verbs may be violated in two ways. First, by insertion of focus-related elements between the nominal and the verb. Second, by scrambling of the nominal to the left and right peripheries of the sentence.

In the former case, a focus particle ham (‘also’) (12a), a contrastive focal clitic –chi (12b) and focused adverbs (12c) may come between the bare nominal and the main verb.

(12) a. Anvar rasm ham chizdi
    Anvar picture also draw.PST.3SG
    ‘Anvar also drew (a) picture(s).’
   b. Anvar kitob-chi, o’qidimi?
    Anvar book-PR read.PST.3SG.Q
    ‘What about (a) book(s), has Anvar read it (them)?’
   c. Anvar kuzda palto kamdan-kam/ ozgina bo’lsa-da/ jinday kiydi
    Anvar autumn.LOC coat rarely briefly wear.PST.3SG
    ‘Anvar wore (a) coat(s) rarely/ a little/ briefly in autumn.’

In the latter case, adjacency-violation is achieved via scrambling of the bare nominal. As shown in (13a), the bare nominal can be dislocated to the left periphery of the sentence, when it serves as a contrastive topic, marking the presence of other alternatives (note that kitob is followed by the contrastive topic marker esa). As shown in (13b), the bare nominal can also undergo rightward scrambling to the post-verbal position, i.e. may be backgrounded.

(13) (based on Öztürk, 2009: 339 for Turkish and Modarresi, 2014: 18 for Persian):
   a. Q: Kim rasm chizdi va kim kitob o’qidi?
      Who picture draw.PST.3SG and who book read.PST.3SG
      ‘Who drew (a) picture(s) and who read (a) book(s)?’
   A: Rasm Anvar chizdi, kitob esa Rano o’qidi
      Picture Anvar draw.PST.3SG book C.T.M Rano read.PST.3SG
      ‘It was Anvar, who drew (a) picture(s) and Rano who read (a) book(s).’

---

6 The default word order in Uzbek is SOV, but scrambling is common and results in the following possible orders: SVO, VSO, OVS, VOS and OSV.
b. Voy, (men) chizdim rasm!
Hey I draw.PST.1SG picture
‘Hey, I did draw (a) picture(s).’

It follows from the discussion in section 2, that linear adjacency of the nominal to the verb is a natural outcome of morpho-syntactic coalescence in TI. In turn, its absence is a strong argument against TI. Neither lexical compounding, nor syntactic incorporation via head-movement, nor head-to-head merge analyses of TI can account for the ability of the incorporated nominal to leave the verb-adjacent position and to have a considerable syntactic mobility. Overall, the data so far strongly suggests the absence of an unbreakable morpho-syntactic tie between bare nominals and main verbs.

Turning to the syntactic status of bare nominals as complements of transitive verbs, the example below illustrates that, just like the full-fledged case-marked direct objects, bare nominals block the possibility of an ‘extra object’ nominal (term due to Chung & Ladusaw, 2004):

(14) *Anvar manzara rasm chizdi
Anvar landscape picture draw.PST.3SG
Intended: ‘Anvar picture-drew (a) landscape(s).’

Based on the combination of the morpho-syntactic evidence discussed in this subsection, we conclude that the treatment of constructions containing bare nominals and main verbs as involving TI is infeasible. Our findings are in line with Öztürk (2005, 2009) and Kamali (2015) for Turkish.

3.1.2 Semantic Evidence

Discussing the semantic evidence in favor of PI in Uzbek, we will illustrate that constructions containing bare nominals, unlike those with regular direct objects, manifest a number of semantic hallmarks of pseudo-incorporation, such as (i) obligatory narrow scope, (ii) number-neutrality, (ii) atypical discourse anaphora of the nominal, and (iv) name-worthiness of the whole verbal construction. Coupled with the morpho-syntactic features discussed in the previous section, the presence of these semantic features strongly suggests that Uzbek employs pseudo-incorporation.

a) Obligatory narrow scope

One of the steadfast cross-linguistic properties of incorporated nominals is ‘obligatory narrow scope’ interpretation, resulting from their scopal inertness. In Uzbek, bare nominals cannot scope over intensional operators (15), negation and universal quantification.

(15) Ra’no maqola o’qishi kerak
Rano article read.3SG must
‘Rano must read (an) article(s).’
Must > ∃x: Rano must read some article(s) or other.
*∃x > Must: There exist(s) (an) article(s) such that Rano must read it (them).
In contrast, non-incorporated regular indefinite (16a) and definite (16b) direct objects allow for a wide scope interpretation.

(16) a. Ra’no bitta maqola(-ni) o’qishi kerak
   Rano one article-ACC read.3SG must
   ‘Rano must read an article.’

   Must > ∃x: Rano must read some article or other.
   ∃x > Must: There exists an article such that Rano must read it.

b. Ra’no maqola-ni o’qishi kerak
   Rano article-ACC read.3SG must
   ‘Rano must read the article.’

   ∃x Article(x): There is a unique article that Rano must read.
   (*Must > ∃x: Rano must read some article or other.)

b) Number-neutrality
Uzbek bare nominals are number-neutral, i.e. although formally singular, such nominals are compatible with both singular and plural interpretation (17a), which is yet another feature of incorporation. In contrast, singular direct objects (i.e. objects with an article and/or case-marking) express strict semantic singularity (17b) and formally plural direct objects, strict semantic plurality (17c).

(17) a. Ra’no maqola o’qidi
   Rano article read.PST.3SG
   ‘Rano read (an) article(s).’

b. Ra’no bitta maqola(-ni) maqola-ni o’qidi
   Rano one article-ACC article-ACC read.PST.3SG
   ‘Rano read an article/ the article.’

c. Ra’no maqola-lar(-ni) o’qidi
   Rano article-PL-ACC read.PST.3SG
   ‘Rano read articles/ the articles.’

c) Atypical anaphora
Turning to the discourse properties of nominals, it is important to note briefly that with the exception of such languages as West Greenlandic and Mapudungun (cf. Sadock, 1980; Bittner, 1994; van Geenhoven, 1998; Baker, 2006), TI languages are characterized by ‘discourse opacity’ of incorporated nominals; i.e. their inability to serve as an antecedent to pronouns in discourse. In contrast, incorporated nominals in most PI languages tend to exhibit different levels of accessibility to various types of anaphora (mostly covert, but also overt) (cf. Farkas & de Swart, 2003; Yanovich, 2008; Dayal, 2011; Espinal & McNally, 2011; Modarresi, 2014; Kamali, 2015; Krifka & Modarresi, 2016).

In Uzbek, similar to other PI languages, bare nominals are not discourse opaque. Moreover, based on their accessibility for both overt and covert anaphoric uptake, we conclude that such nominals are in fact discourse transparent. However, one of the striking differences of anaphoric uptake of bare objects is their unique ability to antecede both singular and plural overt anaphora.

The example below shows that bare nominals are accessible for overt pronouns (18a), the possessive pronominal clitic (18b), as well as covert or null anaphora (18c). Note that the former
two are sensitive to number specification and carry morphological signs of number agreement. Yet, the bare nominal is accessible for both singular and plural anaphoric expressions. The data is well-matched with the number-neutrality of such nominals and with the reported discourse transparency of bare objects in Turkish (Bliss, 2004; Kamali, 2015) and Tatar (Lyutikova & Pereltsvaig, 2015).

(18) Anvar rasm-i chizdi
Anvar picture draw.PST.3SG
‘Anvar drew (a) picture(s).’

a. Men u-ni/ ular-ni tezda sotdim (overt pronominal anaphora)
I it-ACC them-ACC quickly sold.PST.1SG
‘I sold it/ them quickly.’

b. Narx-i/ narx-lar-i ancha baland (possessive pronominal anaphora)
price-POSS.3SG price-PL-POSS.3SG quite high
‘Its price is quite high. / Their prices are quite high.’

c. Ra’no Ø sotib oldi (covert anaphora)
Rano pro buy.CVB take.PST.3SG
‘Rano bought it/ them.’

Compare this to the discourse properties of full-fledged direct objects. As shown in (19), morpho-syntactically singular standard objects can only be picked up by singular anaphoric expressions:

(19) Anvar bitta rasm(-ni)/ rasm-ni chizdi
Anvar one picture-ACC picture-ACC draw.PST.3SG
‘Anvar drew a picture/ the picture.’

a. Men u-ni/ *ular-ni tezda sotdim (overt pronominal anaphora)
I it-ACC/ *it-PL-ACC quickly sell.PST.1SG
‘I sold it quickly.’ (not ‘I sold them quickly.’)

b. Narx-i/ *narx-lar-i ancha baland (possessive pronominal anaphora)
price-POSS.3SG price-PL-POSS.3SG quite high
‘Its price is quite high.’ (not ‘Their prices are quite high.’)

In turn, morpho-syntactically plural objects license plural anaphora only:

(20) Anvar rasm-lar(-ni) chizdi
Anvar picture-PL-ACC draw.PST.3SG
‘Anvar drew (the) pictures.’

a. Men u-lar-ni/ *u-ni tezda sotdim (overt pronominal anaphora)
I it-PL-ACC it-ACC quickly sell.PST.1SG
‘I sold them quickly.’ (not ‘I sold it quickly.’)

b. Narx-lar-i/ *narx-i ancha baland (possessive pronominal anaphora)
price-PL-POSS.3SG price-POSS.3SG quite high
‘Their prices are quite high.’ (not ‘Its price is quite high.’)
Covert anaphora lacks any number sensitivity and, therefore, can be used in relation to all types of full-fledged objects, independently of their number. Turning back to bare nominals, it is important to note that their anaphoric uptake is not unimpeded. Generally, while covert anaphora in relation to bare objects seems unproblematic, speakers’ judgments of overt anaphoric uptake tend to vary. Thus, in some contexts, it seems that one, but not the other, overt anaphoric expression is acceptable. In this regard, Uzbek bare nominals behave similarly to pseudo-incorporated nominals in Persian, whose anaphoric accessibility is determined by world knowledge (cf. Modarresi, 2014; Krifka & Modarresi, 2015).

d) Name-worthiness

The semantic property of name-worthiness in PI concerns the interpretation of the whole incorporated construction. Dayal (2015) uses ‘name-worthiness’ as a general term consolidating a number of interdependent qualities of incorporated units. These qualities include: ‘institutionalization’ (i.e. labeling a recognizable, culturally familiar or habitual activity); ‘gaps’ (i.e. the ungrammaticality of some nominal-verb combinations); and ‘non-compositionality’ (i.e. interpretation of incorporated structures idiomatically).

In Uzbek, constructions containing bare nominals and main verbs seem to be mostly sensitive to the former two qualities, namely institutionalization and gaps. Institutionalization effect captures the interpretation of verbal units as stereotypical or conventional activities. Consider (21), where book-reading is recognized as an institutionalized activity, but word-reading strikes as an odd combination.

(21) a. Anvar kitob o’qidi
    Anvar book read.PST.3SG
    ‘Anvar read (a) book(s).’

b. ??Anvar so’z o’qidi
    Anvar word read.PST.3SG
    ‘Anvar read (a) word(s).’

Institutionalization is frequently responsible for the acceptability of certain combinations that we do not usually expect. In particular, we refer to the potential pseudo-incorporation of animate nominals (22a) and proper names (22b). Although, within the Uzbek DOM paradigm, [+human] objects are most often case-marked, while proper names are obligatorily case-marked, in certain institutionalized activities such nominals may appear bare.

(22) a. Bu shaxs odam o’ldirdi
    This individual human kill.PST.3SG
    ‘This individual murdered (a) human(s).’

b. Biz Samarkand aylandik
    We Samarkand stroll.PST.1PL
    ‘We strolled around Samarkand.’

Concerning the existence of occasional mishaps, gaps, in nominal-verb combinations as in (21b), these are unsystematic and may be bridged. As the cross-linguistic literature on incorporation indicates, such combinations become possible if a context is construed, where the
activity is seen as frequently performed and conventional (see among others, Mithun, 1984; Dayal, 2015 for Hindi; Modarresi, 2014 for Persian; Farkas & de Swart, 2003 for Hungarian). Similarly, Uzbek combinations may be forced to become acceptable if the context is manipulated accordingly. For instance, if we imagine that a certain competition involves a task of word-reading, then this combination becomes acceptable in a given context.

Although, idiomatization is not a characterizing quality of constructions containing bare nominals and main verbs, there seems to be a form of semantic expansion of the meaning of the whole construction. In (23a) the combination tea-drink may be interpreted quite broadly as involving other activities, such as having a snack or a treat, as well as social interaction. In turn, with a parallel definite object, the meaning of the combination is always compositional (23b).

(23) a. Kel, choy ichamiz
    ‘Come, we will drink tea.’

b. Kel, choy-ni ichamiz
    ‘Come, we will drink the tea.’

To sum up, the morpho-syntactic and semantic evidence presented in subsections 3.1.1 and 3.1.2, respectively, strongly indicates that bare objects of main verbs undergo a type of object incorporation that must be analyzed as pseudo-incorporation, rather than true incorporation.

3.1.3 Analysis of Pseudo-Incorporation in Uzbek

We follow the general trend in the research of treating incorporated nominals as denoting properties, i.e. of the type \(<e,t>\) (cf. van Geenhoven 1996, 1998; Dayal, 2003, 2011, 2015; Farkas & de Swart, 2003; Chung & Ladusaw, 2004; Kagan, 2005, 2012, 2015; Dobrovie-Sorin et al., 2006; Espinal & McNally, 2011; Modarressi, 2014). A range of proposals have been made in the literature regarding the semantic combination of a property-denoting bare nominal and a transitive verb of the type \(<e,<e,t>>\) (cf. references listed above.) The Uzbek data discussed above is compatible with various accounts. One option is to adopt for Uzbek Chung & Ladusaw’s (2004) analysis and its adaptation by Modarressi (2014) for Persian PI. Under this account, in Uzbek PI, the meaning of the predicate gets restricted to the meaning of the nominal via RESTRICT; i.e. an operation that facilitates a direct combination of a property and a predicate. The property in such a combination is a restrictive modifier. Since predicates are functions, the initial domain of the predicate gets restricted to its subdomain. The application of RESTRICT is illustrated in (24a), where, as the result of the operation, the meaning of the predicate drew gets restricted to the meaning of the nominal picture, with the latter acting as a restrictive modifier of the former.

(24) Anvar rasm chizdi
    ‘Anvar drew (a) picture(s).’

a. RESTRICT (λyλx [draw’(y)(x)], picture’) = λyλx [draw’(y)(x) ∧ picture’(y)]
b. EC (RESTRICT (λyλx [draw’(y)(x)], picture’)) = λx∃y [draw’(y)(x) ∧ picture’(y)]
RESTRICT is a non-saturating mode of composition: when a transitive predicate of the type 
\(<e,<e,t>>\) combines with a bare nominal of the type \(<e,t>\), the latter does not satisfy the argument position of the former. Therefore, Chung & Ladusaw (2004) propose that the total semantic predicate saturation takes place via existential closure (EC). This operation occurs immediately above the verbal phrase (vP) and provides an existential binding to the argument variable (24b). Consequently, EC closes off the predicate’s argument position.

This analysis accounts for a range of special properties that characterize Uzbek bare nominals. Firstly, their ‘deficient morpho-syntax’ is captured by the proposal that these are bare NPs. Their inability to contain demonstratives, quantifiers, plural morphology and other functional elements is related to the absence of the NumP and DP projections in the structure. (Obviously, an object may contain a numeral, a demonstrative, etc., but in this instance, it is no longer bare and does not exhibit the set of PI properties described above.) Second, semantic number neutrality is, again, linked to the absence of the NumP projection. We assume that morphologically, the nominals are singular since the singular constitutes the default form of Uzbek nouns (whereas in order to create the plural, the suffix \(-lar\) has to be added).

Secondly, the proposed analysis explains the scopal behavior of PI objects, namely their ‘obligatory narrow scope’ interpretation. Property-denoting expressions are expected to be scopally inert. Further, bare nominals reside below the existential closure. This, too, explains why Uzbek bare nominals cannot receive wide scope interpretation. Note that although it has been shown that the bare nominals (unlike truly incorporated ones) may occur in a non-verb-adjacent manner, they still obligatorily receive the narrow scope reading as shown in (25). This, again, indicates that PI nominals are interpreted below EC.

\[(25)\]
\[\text{Rasm, Anvar chizishi kerak}\]
\[\text{Picture Anvar draw.3SG must}\]
\[\text{‘It is Anvar who must draw (a) picture(s).’}\]
\[Must > \exists x: \text{Anvar must draw some picture(s) or other(s).}\]
\[*\exists x > Must: \text{There exist(s) (a) picture(s) such that Anvar must draw it (them).}\]

Thirdly, the current analysis accounts for the ‘name-worthiness’ of the whole verbal construction. In particular, the result of predicate restriction is that the domain of the predicate serving as a function gets narrowed down to its subdomain, e.g. from ‘drew’ to ‘picture-drew’. The nominal in such a construction is number-neutral, denotes a property and does not refer to a particular or a specific object, serving as a mere restrictive modifier of the verb. As a result, the construction itself tends to be interpreted as a unitary concept, used as a label for a recognizable, culturally familiar, or habitual activity; i.e. an institutionalized activity. Note also that the possibility of forcing gaps into becoming institutionalized activities follows from the identified process behind predicate restriction. Specifically, it indicates that, since the nominal denotes a property, it can potentially enter into the RESTRICT relation with the verb.

Finally, the ‘atypical discourse anaphora’ of Uzbek bare objects is particularly difficult to account for, given that the descriptive facts are complex and subject to variation in native speaker judgments. One possibility is that the pseudo-incorporated nominals do not create a discourse referent in the strict sense of the term. Rather, anaphoric expressions get their reference through (bridging) inference (cf. e.g. Clark 1977, Prince 1981). This would account for the fact that, depending on the context, the anaphor may be singular or plural. Also, different speakers may find this kind of inference easier or more difficult, which results in variation in judgments.
Alternatively, the mechanism proposed by Krifka & Modarresi (2016) can be adopted, whereby PI bare nominals introduce discourse referents, but (unlike standard non-incorporated objects) do not do so straightforwardly; rather, the relative transparency of bare objects results from certain manipulations in the DRS. The Abstraction & Summation rule is introduced which enables the anaphoric uptake of the discourse referent. For reasons of space we do not expand on this direction here, but see Levy-Forsythe (2018) for a detailed account.

3.2 True Incorporation

The examination of bare nominal+verb constructions reveals that not all of them behave uniformly. Besides main verbs discussed in the previous section, Uzbek has a more limited class of light verbs (term due to Jespersen, 1954), i.e. verbs with a weak or bleached semantic content. Cross-linguistically, they form what is known as light verb constructions (LVCs), where the meaning of the verb is completed by a nominal (Grimshaw & Mester, 1988; Choi & Wechsler, 2001; Kearns, 2002; Butt, 2003; Karimi-Doostan, 2005; Megerdoomian, 2012). In Uzbek, light verb constructions containing a bare nominal and a light verb seem to display a tight morpho-syntactic interaction. Bare nominals in them (i) have a minimal structure and (ii) exhibit true morpho-syntactic fusion with the verb, which may affect verbal valency. At the same time, light verb constructions align with the semantic properties of TI: (i) obligatory narrow scope, (ii) number-neutrality, (ii) discourse opacity of the nominal and (iv) strong name-worthiness effect. Below, we propose that the combination of the syntactic and semantic characteristics serves as evidence against their PI and in favor of TI.

3.2.1 Morpho-Syntactic Evidence

In Uzbek, by light verbs we refer to semantically weak verbs such as qil- and et- (‘do’, ‘make’), ol- (‘take’) or ber- (‘give’), as well as verbs that appear with bare nominals in idiomatic expressions, such as go’l go’y- (lit. ‘hand-put’ – ‘sign’), guloq sol- (lit. ‘ear-put’ – ‘listen’), ko’z urish tir- (lit. ‘eye-clash’ – ‘flirt’), etc. Bare nominals in LVCs are deficient in their functional architecture and syntactic status in a clause, compared to both regular and PI objects.

a) Minimal nominal structure

Bare nominals in LVCs are completely stripped of any functional layers, i.e. have a minimal nominal structure. They disallow direct adjectival (26a) or RC (26b) modification. As indicated in the translation of (26a), modification concerns the whole incorporating unit and not the bare nominals itself (compare to (8a)). In turn, RC modification is ungrammatical.

(26) a. Anvar g’oyibona rasm soli
   Anvar secret(ly)  picture put.PST.3SG
   1. ‘Anvar drew (a) picture(s) secretly,’ (= ‘Anvar drew secretly.’)
   2. #‘Anvar drew (a) secret picture(s).’

   b. *Anvar hammani hayratlan tir-adi-gan  rasm sol-adi
   Anvar everyone.ACC astonish-CAUS-IMPRF.3SG-PTCPL picture put.IMPRF.3SG
   Intended: ‘Anvar draws (a) picture(s) that astonish(es) everyone.’
Besides, bare plural nominals are disallowed as complements of light verbs, additionally indicating that the incorporated nominal in LVC is completely bare:

(27) *Ra’no qo’ng’iroq-lar qildi
    Rano call-PL make.PST.3SG
    Intended: ‘Rano made phone calls.’

We conclude that bare nominals in LVCs are bare N₀ heads.

b) Syntactic rigidity
The relation between bare nominals and light verbs is tight, as expected between the elements of a TI unit. For instance, a bare nominal in a LVC may not appear as a stand-alone answer to a ‘wh’-question, which indicates that it is not a syntactic constituent:

(28) *Q: Anvar nima soldi?             A: *Rasm
    Anvar what put.PST.3SG picture
    Intended: ‘Q: What did Anvar put? A: (a) picture(s).’

Further, in many instances, it is impossible to omit the light verb under identity (29a), conjoin verbs serving as predicates of the same nominal (29b) and conjoin bare nominals in the preverbal position (29c).

(29) (adapted from Öztürk for Turkish, 2009: 339):
   a. *Anvar rasm emas, quloq soldi
      Anvar picture not ear put.PST.3SG
      Intended: ‘Anvar did not picture-put, he ear-put.’ (= ‘Anvar didn’t draw, he listened.’)
   b. *Anvar rasm soldi va menga ko’rsatdi
      Anvar picture put.PST.3SG and I.DAT show.PST.3SG
      Intended: ‘Anvar drew (a) picture(s) and showed it (them) to me.’
   c. *Anvar rasm va quloq soldi
      Anvar picture and ear put.PST.3SG
      Intended: ‘Anvar picture- and ear-put.’ (= ‘Anvar drew and listened.’)

In addition, bare nominals in LVCs are strictly verb-adjacent. They may not be split from the verb by the focus particle ham (‘also’) (30a)⁷, the contrastive focal clitic -chi (30b) or focused adverbs of degree (30c):

(30) a. ??Anvar rasm ham soldi
      Anvar picture also put.PST.3SG
      Intended: ‘Anvar also drew (a) picture(s).’ (= ‘Anvar also drew.’)
   b. *Anvar qo’ng’iroq-chi, qildimi?

---

⁷ Under the reading when the whole LVC (and not just a bare nominal) is focused, the insertion of ham may be possible. This is parallel to the Tatar focus particle –gına reported by Lyutikiva & Pereltsvaig (2015: 309), which attaches to bare nominals within LVCs, but must be interpreted as focusing the entire verbal construction.
Anvar call-PR make.PST.3SG.Q
Intended: ‘What about (a) phone call(s), did Anvar make it (them)?’
c. *Anvar kuzda sayr kamdan-kam/ ozgina bo‘lsa-da/ jinday etdi
Anvar autumn.LOC stroll rarely a little although briefly make.PST.3SG
Intended: ‘Anvar took (a) stroll(s) rarely/ a little/ briefly in autumn.’

Moreover, such bare nominals exhibit syntactic immobility: they may not be dislocated from
the preverbal position to the left and right peripheries of the sentence.

(31) a. Q: Kim g‘iybat qildi va kim quloq osdi?
Who gossip make.PST.3SG and who ear hang.PST.3SG
‘Who gossiped and who listened?’
   A: *G‘iybat Anvar qildi, quloq esa Ra’no osdi
      Gossip Anvar make.PST.3SG ear C.T.M Rano hang.PST.3SG
      Intended: ‘It was Anvar, who gossiped and Rano who listened.’
b. *Voy, (men) soldim rasm!
Hey I put.PST.1SG picture
Intended: ‘Hey, I did draw (a) picture(s).’ (= ‘Hey, I did draw.’)

Finally, nominal components of LVCs may lead to a change of verb’s valency. Informally
speaking, ‘object-doubling’ is allowed (cf. ex. (14)). Thus, some LVCs containing bare nominals
may take both a full-fledged accusative-marked direct object and a pseudo-incorporated bare
nominal (32a), while some may require a dative-marked indirect object (32b).

(32) a. Anvar manzara(-ni) tasvir etdi
      Anvar landscape-ACC depiction make.PST.3SG
      ‘Anvar depicted (a) landscape(s).’/ ‘Anvar depicted the landscape.’
b. Anvar rasm-ga egalik etdi
      Anvar picture-DAT ownership make.PST.3SG
      ‘Anvar owned a picture.’

To sum up thus far, bare nominals in LVCs exhibit syntactic behavior that is characteristic of
bare Ns that undergo TI.

3.2.2 Semantic Evidence

Semantically, these nominals, again, exhibit properties that are characteristic of incorporation.

a) Obligatory narrow scope
Like bare nominal complements of main verbs, bare nominals in LVCs are indefinite and non-
specific and do not allow a wide scope interpretation.

(33) a. Anvar mashq qilishi kerak
      Anvar exercise do.3SG must
      ‘Anvar must exercise.’
Must > ∃x: Anvar must do some exercise(s) or other.
*∃x > Must: There exist(s) (an) exercise(s) such that Anvar must do it (them).

b) **Number-neutrality**
Bare nominals in LVCs are number-neutral. Thus, (34a) could mean that Anvar did one or more exercises (or more naturally, that Anvar engaged in exercising in general).

(34) Anvar mashq qildi
    Anvar exercise do.PST.3SG
    ‘Anvar did (an) exercise(s).’ / ‘Anvar exercised.’

c) **Discourse opacity (Atypical anaphora)**
Nominals in LVCs are discourse opaque, i.e. inaccessible for anaphoric uptake (which is typical of TI but not necessary at all in PI).

(35) Anvar rasm$_i$ soldi
    Anvar picture put.PST.3SG
    ‘Anvar drew (a) picture(s).’ (= ‘Anvar drew.’)
   a. *Men u-ni/ ular-ni tezda sotdim (overt pronominal anaphora)
      I it-ACC them-ACC quickly sold.PST.1SG
      Intended: ‘I sold it/ them quickly.’
   b. *Narx-i/ narx-lar-i ancha baland (possessive pronominal anaphora)
      Price-POSS.3SG price-PL-POSS.3SG quite high
      Intended: ‘Its price is quite high.’ / Their prices are quite high.’
   c. *Ra’no Ø sotib oldi (covert anaphora)
      Rano pro buy.CVB take.PST.3SG
      Intended: ‘Rano bought it (them).’

d) **Name-worthiness**
Name-worthiness is a full-blown characteristic of LVCs. Compared to pseudo-incorporated constructions, LVCs show stronger institutionalization, stricter gaps and high levels of idiomatization.

For instance, the interpretation of the LVC ko’z yumdi (lit. ‘eye-closed’) in (36a) is idiomatic and refers to passing away, thus denoting an institutionalized activity. In contrast, the regular direct object constructions in (36b) lack an institutional reading and must be interpreted non-compositionally.

(36) a. Shoir ko’z yumdi
    Poet eye close.PST.3SG
    ‘The poet passed away.’
   b. Shoir bitta ko’z(-ni)/ ko’z-ni yumdi
    Poet one eye-ACC eye-ACC close.PST.3SG
    ‘The poet closed an eye/ the eye.’ (not ‘The poet passed away.’)

Generally, gaps in LVCs seem to occur more often and are harder to remedy by contextual manipulations, indicating a higher level of lexicalization. Consider the following examples,
where the first category of combinations is grammatical (37a), and the second category, although quite close in meaning, is ungrammatical (37b).

(37) (adapted from Dayal, 2011: 133):

a. mehnat qilmoq; bola ko‘rmoq; qiz uzatmoq; qovun tushirmoq
   labor-make  child-see  daughter-pass  melon-drop
   ‘work’      ‘give birth’ ‘marrying off (a) daughter(s)’ ‘make (a) blunder(s)’

b. *mashaqqat qilmoq; *ayol ko‘rmoq; *o‘g‘il uzatmoq; *tarvuz tushirmoq
   hardship-make  woman-see  son-pass  watermelon-drop

Based on the syntactic and semantic facts discussed above we propose that bare nominals in light verb constructions undergo true incorporation in Uzbek. In other words, this language exhibits both TI and PI, depending on the type of verb involved. We leave for future research the question of whether the incorporation takes place in the lexicon or in the syntax.

4 Conclusion

Uzbek facts reveal that a given language may exhibit instances of both true incorporation and pseudo-incorporation, depending on the type of the verb involved.

- Bare nominal components of LVCs are truly incorporated.

  The “object” is an N⁰ head which does not project its own phrase; rather, the N-V string is treated by the syntax as a V head.

- Bare nominal complements of other, main, verbs are pseudo-incorporated.

  The object is a NP which is merged in the complement position and combines with the V to form a VP. Semantically, we follow Chung & Ladusaw’s (2004) approach, whereby a property-denoting bare nominal and the verbal predicate are combined via the RESTRICT function.
References


A NOTE ON LEXICAL RECIPROCITY IN BRAZILIAN PORTUGUESE

GIADA PALMIERI
RENATO MIGUEL BASSO
MANUELA PINTO
YOAD WINTER
JOOST ZWARTS

Department of English Studies, Utrecht University
Department of Linguistics, Universidade Federal de São Carlos

1 Introduction

In the past decades, extensive research has focused on the way in which languages express reciprocal configurations (Dalrymple et al., 1998; Nedjalkov, 2007; König & Gast, 2008, a.o.). As for verbal reciprocals, two strategies are available cross-linguistically. One strategy, grammatical reciprocity, consists in the use of a grammatical element, like a reciprocal pronoun, to generate a mutual configuration from any transitive verb, as shown by the English example in (1). On the other hand, there are verbs that convey a reciprocal interpretation using their intransitive entry: this is the case for verbs that undergo the reciprocal alternation (Levin, 1993), also referred to as lexical reciprocals. These verbs constitute a restricted class, which typological works characterize as rather stable cross-linguistically (Haspelmath, 2007).

(1) Irene and Sara hugged/ kissed/ thanked/ punished each other

(2) Irene and Sara hugged/ kissed/ *thanked/ *punished

Work by Palmieri, Winter and Zwarts was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 742204).
English makes an overt distinction between grammatical and lexical reciprocal forms, and parallel surface distinctions are available in other languages, such as Hebrew, Dutch and Arabic (Reinhart & Siloni, 2005). However, the identification of lexical reciprocal predicates is not straightforward in a number of languages. In Romance, for instance, only one reciprocal form is generally available in finite clauses, where the presence of the clitic si/se (in its various forms) allows a reciprocal interpretation with any transitive verb (3). Note that the same construction also gives rise to a reflexive interpretation, but this will not be included in our discussion.¹

(3) a. Irene e Sara si abbracciano/ puniscono  Italian
    b. Irene y Sara se abrazan/ castigan  Spanish
    c. Irene et Sara s’embrassent/ punissent  French

Irene and Sara hug/PAST.PRES.3P punish/PRES.3P
i. ‘Irene and Sara hug/punish each other’
ii. ‘Irene and Sara hug/punish themselves’

Despite the existence of only one surface reciprocal form, it has been shown that verbs with lexical reciprocal entries do exist in Romance languages. Through systematic tests, it is possible to identify a group of verbs that show the same semantic and morpho-syntactic properties that characterize lexical reciprocals cross-linguistically (see Palmieri et al. 2018 on lexical reciprocals in Italian, and Doron & Rappaport Hovav 2009 on lexical reflexives in French).

In most Romance languages, the clitic si/se is obligatory in all simple reciprocal sentences. However, Brazilian Portuguese (henceforth BrP) represents an exceptional case: for some verbs in this language, reciprocal interpretations are available with or without se (4a). Most verbs, nonetheless, do require se (4b), similarly to the conventional Romance pattern in (3).

(4) a. Irene e Sara (se) abraçaram/ beijaram  Brazilian Portuguese
   Irene and Sara hug/PAST.PRES.3P kiss/PAST.3P
   ‘Irene and Sara hugged/kissed (each other)’
   b. Irene e Sara *(se) puniram/ agradeceram  Brazilian Portuguese
   Irene and Sara punish/PAST.3P thank/PAST.3P
   ‘Irene and Sara punished/thanked each other’

The distribution of se in BrP is intriguing, not only because of the surprising difference from other Romance languages, but also because of the resemblance to languages like English, where verbs like hug and kiss give rise to a reciprocal interpretation in their intransitive entry. We hypothesize that BrP verbs that support reciprocal interpretations without se have a lexical reciprocal entry. To examine this hypothesis, we will look at semantic properties that have been cross-linguistically associated with lexical reciprocity, and used to identify verbs with a lexical reciprocal entry in Italian (Palmieri et al. 2018). Subsequently, we will investigate to what extent the meanings of reciprocal verbs identified in BrP correspond to those of the Italian verbs that we identified in Palmieri et al. (2018).

We propose that the BrP verbs that allow the omission of se in finite clauses have a lexical reciprocal entry, as they show semantic properties that are cross-linguistically associated with lexical reciprocity: (i) they can appear in the discontinuous reciprocal construction and (ii) allow

¹ Abbreviations used in this paper: PRES = present tense; INF = infinitive; AUX = auxiliary; PP = past participle; PAST= simple past tense; 1S = 1st person singular; 3S = 3rd person singular; 3P= 3rd person plural; si = clitic si/se.
A note on lexical reciprocity in Brazilian Portuguese

A reciprocal interpretation with singular group NPs. A comparison with Italian will reveal a great correspondence between the lexical reciprocals in these two languages. Furthermore, we propose that semantically, Italian and BrP do not differ in the way they encode reciprocal meanings, despite the evident difference in finite clauses. We propose that the omission of se in BrP finite clauses is parallel to the omission of si in Italian non-finite complemental clauses, which does not prevent lexical reciprocal verbs from conveying a reciprocal interpretation (Palmieri et al. 2018). We conclude that Italian and BrP have different syntactic parameters regulating the obligatory use of si/se in finite clauses, but that this element cannot be considered as the source of lexical reciprocity in either of these languages.

The paper is structured as follows. In §2 we focus on the identification of lexical reciprocity in BrP: we illustrate which verbs generate a reciprocal interpretation without se, and spell out semantic properties that characterize these verbs, analogously to lexical reciprocals in other languages. In §3 we will compare the class of lexical reciprocal verbs in BrP to those of Italian, unveiling a strong correspondence between the concepts that are lexicalized as reciprocal in the two languages. In §4 we will briefly discuss grammatical reciprocity in BrP. In §5 we provide general conclusions.

2 Lexical Reciprocity in BrP

In the rich literature on the Romance clitic si/se, it has often been proposed that this element cannot be analyzed as a pronoun: across Romance languages there are verbs with lexical entries where the reflexive or reciprocal meaning resides in the verb root (Labelle 2008; Doron & Rappaport Hovav 2009; Palmieri et al. 2018). It has been proposed that with such verbs, si/se does not contribute to the reciprocal interpretation, but functions as a marker whose obligatory presence is to be analyzed in terms of syntactic requirements. Evidence in favor of this account comes from constructions where si/se is disallowed but reflexive/reciprocal meanings are still available with a restricted number of verbs. Doron & Rappaport Hovav (2009) showed that some French verbs, if embedded in a causative construction where se is not allowed for independent reasons, would nonetheless allow a reflexive interpretation. A parallel behavior has been observed in Italian lexical reciprocals (Palmieri et al. 2018): verbs like abbracciare ‘to hug’ and baciare ‘to kiss’ obligatorily require si in finite clauses (5a). However, in causative clauses (5b), these verbs are distinguished from other transitive verbs in supporting reciprocal interpretations without any grammatical marking.

\[
\begin{align*}
\text{a. Irene} & \quad \text{Sara si abbracciano/ baciano} \\
& \quad \text{Irene and Sara si hug.PRES.3P kiss.PRES.3P} \\
& \quad \text{‘Irene and Sara hug/kiss (each other)’} \\
\text{b. Ho fatto} & \quad (*\text{si}) \text{ abbracciare / baciare Irene e Sara} \\
& \quad \text{have.AUX.1SG make.PP si hug.INF kiss.INF Irene and Sara} \\
& \quad \text{‘I caused Irene and Sara to hug/kiss’}
\end{align*}
\]

Along these lines, it seems plausible that Romance languages systematically express reciprocal meanings similar to ‘hug’ and ‘kiss’ without si/se. Moreover, if the analysis of si/se as a marker is on the right track, and its obligatory use in finite clauses is linked to reasons that are external to the semantics of lexical reciprocals, the omission of si/se could in principle also be possible in
constructions other than causatives. We propose that this is the case with finite clauses in BrP, where *se* can be omitted with lexically reciprocal verbs. In order to provide support for this proposal, let us start by illustrating the distribution of *se* in BrP.

### 2.1 Reciprocal Meanings without *se*

In BrP, the presence of *se* is not always a precondition for a reciprocal interpretation in finite clauses. The possibility to omit *se* depends on the verb. In most cases, *se* is required when describing a reflexive or reciprocal situation, hence omitting this element results in ungrammaticality (6). However, there are verbs that disallow *se*, yet get a reciprocal interpretation without it (7a), and verbs for which the presence of *se* is optional in reciprocal sentences (7b-c). When sentences with these latter verbs omit *se*, they only receive a reciprocal interpretation, ruling out the reflexive reading.²

#### (6) Irene e Sara *(se) puniram/* agradeceram

Irene and Sara SI punish.PAST.3P thank.PAST.3P

i. ‘Irene and Sara punished/thanked each other’

ii. ‘Irene and Sara punished/thanked themselves’

#### (7) a. Irene e Sara *(se) discutiram/* colaboraram

Irene and Sara SI discuss.PAST.3P collaborate.PAST.3P

‘Irene and Sara discussed/collaborated’

b. Irene e Sara se abraçaram/* beijaram

Irene and Sara SI hug.PAST.3P kiss.PAST.3P

i. ‘Irene and Sara hugged/kissed each other’

ii. ‘Irene and Sara hugged/kissed themselves’

c. Irene e Sara abraçaram/* beijaram

Irene and Sara hug.PAST.3P kiss.PAST.3P

‘Irene and Sara hugged/kissed each other’

While the behavior in (6) is similar with most transitive verbs, the obligatory or optional omission of *se* is a characteristic of a limited number of verbs, displayed respectively in Table 1 and 2.

---

² There are also verbs that can optionally omit *se* and generate a reflexive interpretation (i)-(ii). We assume that these verbs have a lexical reflexive entry, but they will not be included in our discussion.

#### (i) Irene e Sara se depilararam

Irene and Sara SI depilate.PAST.3P

i. ‘Irene and Sara depilated themselves’

ii. ‘Irene and Sara depilated each other’

#### (ii) Irene e Sara depilaram

Irene and Sara SI depilate.PAST.3P

‘Irene and Sara depilated themselves’
A note on lexical reciprocity in Brazilian Portuguese

BrP verbs that rule out se in finite clauses, and support reciprocal readings without se

<table>
<thead>
<tr>
<th>BrP verbs that support reciprocal readings in finite clauses with optional se</th>
</tr>
</thead>
</table>

Table 1. BrP verbs that do not allow se in finite clauses, as in (6)

<table>
<thead>
<tr>
<th>BrP verbs that optionally take se in finite clauses, as in (7)</th>
</tr>
</thead>
</table>

Table 2. BrP verbs that optionally take se in finite clauses, as in (7)

The predicates listed above exhibit some similarities between their binary forms. As illustrated in (8), the verbs in Table 1 can introduce another reciprocal argument using the preposition com ‘with’ (i.e. they undergo the ‘with’ simple reciprocal alternation, Levin 1993). Some of the verbs in Table 2 show the same pattern (9a), while some of them can take a direct object (i.e. they undergo the simple transitive reciprocal alternation, in Levin’s terminology), but can also optionally use com ‘with’ to introduce the object, as in (9b).

(8) Irene discuss.PAST.3S colaborou com a Sara
Irene discussed/collaborated with Sara

(9) a. Irene casou *(com) a Sara
Irene marry.PAST.3S with the Sara
‘Irene married Sara’

b. Irene abraçou (com) a Sara
Irene hug.PAST.3S with the Sara
‘Irene hugged (with) Sara’

Despite this variation, all the verbs in Tables 1 and 2 allow their collective intransitive entry to appear without se (either obligatorily or optionally). This leads us to hypothesize the existence of a lexical reciprocal entry. This assumption does not only rely on finite clauses, but also on the visible correspondence between these verbs and the class of lexical reciprocals surveyed in typological works (Haspelmath, 2007), as well as the Italian lexical reciprocals that give rise to a reciprocal interpretation without si in causatives (Palmieri et al. 2018).³

³ In Italian there is a crucial contrast between finite and causative clauses: in order to generate a reciprocal interpretation, si is obligatory in the former but disallowed in the latter – see example (5). BrP differs in this respect, because se is allowed in causatives, where it has the same distribution as in finite clauses. In causatives, in order to have a reciprocal reading, se is obligatory with verbs that require this element in finite clauses (i), but it can be omitted with verbs that optionally drop se in finite clauses (ii). Thus, given the analogous distribution of BrP se across finite and causative clauses, we will only systematically look at the former.
This intuitive assumption requires further evidence, in order to exclude the possibility that the omission of se in BrP is an idiosyncratic property of certain verbs. Thus, in the next subsections we will provide independent evidence in favor of our hypothesis, by showing that verbs which omit se display two properties that also characterize lexical reciprocals in other languages: they can appear in the discontinuous reciprocal construction (§2.2) and allow a reciprocal interpretation with morpho-syntactically singular NPs (§2.3).

2.2 Discontinuous Reciprocal Construction

The discontinuous reciprocal construction is a construction where one reciprocal argument is encoded as syntactic subject, while another reciprocal argument is introduced by the preposition with. It has been noted that this construction is generally available with lexical reciprocal verbs (Kemmer 1993, Dimitriadis 2004, Siloni 2012), as illustrated in the Hebrew examples below: in (10a) the lexical reciprocal verb hug appears in the discontinuous construction, with an interpretation according to which the boys and the girls hugged. However, the same verb, in its grammatical reciprocal form, leads to ungrammaticality in this construction (10b).

(10) a. ha-yeladim hitnašku im ha-yeladot
the boys kissedREC with the-girls
‘the boys kissed with the girls’
(Siloni 2012, p. 297)
b. *ha-yeladim nišku exad et ha-šeni im ha-yeladot
the boys kiss each other with the girls

In BrP, we have seen that the verbs which do not allow se in finite clauses require com ‘with’ to introduce the second argument. Therefore, their binary entry overlaps with the discontinuous reciprocal construction (11a). Discontinuous reciprocity is furthermore possible with all verbs for which se is optional in finite clauses (i.e. those listed in Table 2), as shown in (11b). Note that for these verbs, the presence of se is also optional in this construction.

(11) a. Eu fiz Irene e Sara se punirem
I make.PP Irene and Sara SI punish.INF
‘I caused Irene and Sara to punish each other’
(b) Eu fiz Irene e Sara (se) abraçarem
I make.PP Irene and Sara (se) hug.INF
‘I caused Irene and Sara to hug’

On the other hand, verbs that require se to express reciprocal configurations in finite clauses, lead to ungrammaticality in the discontinuous reciprocal construction (12a). We found a few exceptions in this respect: the verbs encontrar ‘to meet’, consultar ‘to consult’, falar ‘to talk’, corresponder ‘to correspond’, unir ‘to merge/combine’, sobrepor ‘to overlap’ and confundir ‘to

(i) Eu fiz Irene e Sara se punirem
I make.PP Irene e Sara se Punish
(ii) Eu fiz Irene e Sara (se) abraçarem
I make.PP Irene and Sara (se) abraçarem

4 The only exception in this respect is beijar ‘to kiss’, which takes an optional se in finite clauses, but leads to ungrammaticality in the discontinuous reciprocal construction, regardless of the presence of se:

(i) *Irene (se) beijou com a Sara
Irene SI kiss.PAST.3S with the Sara

5 In this respect, the verbs debate ‘to debate’ and conversar ‘to converse’ constitute an exception, as they do not allow se in this construction, although se is optional in finite clauses.

(i) Irene (*se) debatou/ conversou com a Sara
Irene SI debate.PAST.3S converse.PAST.3S with the Sara

‘Irene debated/ conversed with Sara’
confuse’ can appear in the discontinuous reciprocal construction, although they obligatorily require se (12b).

(11) a. Irene discutiu / colaborou com a Sara
Irene discuss.PAST.3S collaborate.PAST.3S with the Sara
‘Irene discussed/collaborated with Sara’
b. Irene (se) abraçou / casou com a Sara
Irene SI hug.PAST.3S marry.PAST.3S with the Sara
‘Irene hugged/married with Sara’

(12) a. *Irene se puniu / agradeceu com a Sara
Irene SI punish.PAST.3S thank.PAST.3S with the Sara
b. Irene se encontrou / consultou com a Sara
Irene SI meet.PAST.3S consult.PAST.3S with the Sara
‘Irene met/conferred with Sara’

In line with previous works suggesting that the lexical reciprocal entry is a condition for the availability of the discontinuous reciprocal construction, the BrP data above provide evidence in favor of our hypothesis: the verbs that can convey a reciprocal interpretation without se in finite clauses also allow discontinuous reciprocity. The few other verbs that allow discontinuous reciprocity but require se have meanings that are often associated with reciprocal verbs, hence we speculate that their requirement of se is an idiosyncrasy.

2.3 Morpho-syntactically Singular Group NPs

Group nouns are nouns such as committee, team or couple in English, which can take a plural but not a singular of-complement (Barker, 1992):

(13) a. A team/ committee/ couple of women
b. *A team/ committee/ couple of woman

Group NPs constitute an interesting instrument for identifying lexical reciprocity: across different languages, lexical reciprocal verbs lead to considerably different interpretations than grammatical reciprocals when associated with group NPs (Authier & Reed 2018, Palmieri et al. 2018). In English, each other is unacceptable with all verbs and singular group-denoting subjects (14a). In Italian, with most verbs, si only generates a reflexive interpretation with morpho-syntactically singular group NPs: (14b) is interpreted with the team (as a whole) punishing itself. By contrast, both in English and Italian, lexical reciprocal verbs yield a collective interpretation if combined with group NPs: the examples in (15) are interpreted with the members of the team hugging.

6 Note that the verbs unir ‘to merge/combine’, sobrepor ‘to overlap’ and confundir ‘to confuse’ express a reciprocal configuration among the objects of the binary entry (i) and the subjects of the unary entry (ii):

(i) Irene uniu água e farinha
   Irene merge.PAST.3S water and flour
   ‘Irene combined water and flour’
(ii) Água e farinha se uniram
   water and flour SI merge.PAST.3P
   ‘Water and flour combined’
An analogous distinction can be found in BrP, where the availability of a reciprocal interpretation with singular group nouns is also restricted. As exemplified in (16a), the BrP verbs that obligatorily omit se allow a reciprocal interpretation with the group noun a time ‘the team’, in parallel to the reciprocal interpretation of the English and Italian examples in (15). The same reading is available for the verbs with an optional se (16b).

On the other hand, verbs that cannot drop se in finite clauses do not generally allow a reciprocal interpretation with singular group NPs, but rather a reflexive one (17a), similarly to the Italian example in (14b). However, also in this respect, some verbs constitute an exception: encontrar ‘to meet’, consultar ‘to consult’, falar ‘to talk’, corresponder ‘to correspond’, unir ‘to merge/combine’, sobrepor ‘to overlap’ and confundir ‘to confuse’ yield a reciprocal interpretation, although the presence of se is obligatory, both in finite clauses and in (17b). Importantly, these are the same verbs that also allow the discontinuous reciprocal construction.

(16)  
(a) O time (*se) discutiu / colaborou  
the team SI discuss.PAST.3S collaborate.PAST.3S  
‘The members of the team discussed/collaborated’
(b) O time (se) abraçou / separou  
The team SI hug.PAST.3S separate.PAST.3S  
‘The team hugged/separated’

(17)  
(a) #O time se puniu / agradeceu  
the team SI punish.PAST.3S thank.PAST.3S  
‘#The team punished/thanked itself’

7 Similarly to what has been observed with the discontinuous reciprocal construction (see footnote 5), the presence of se across different constructions is not the same as in finite clauses. Among the verbs that have an optional se in finite clauses, three require se in order to give a reciprocal interpretation with group NPs (i), while one verb does not allow se in this environment (ii).

(i) O time *(se) parece/ beija/ misturou  
The team SI resemble.PRES.3S kiss. PRES.3S mix. PRES.3S  
‘The members of the team resemble/kissed/mixed with each other’
(ii) O time *(se) debate  
The team SI debate PRES.3S  
‘The members of the team debate’
A note on lexical reciprocity in Brazilian Portuguese 75

b. O time se encontrou / consultou
the team SI meet.PAST.3S contult.PAST.3S
‘The team met/conferred’

These data provide support in favor of a lexical/grammatical distinction in BrP, based on a semantic property that has been observed in other languages. The division that emerged is in line with our hypothesis: a reciprocal interpretation with singular group NPs is available with verbs that can omit se in finite clauses, and, except for a few exception, the converse is true as well.

3 Lexical Reciprocals in BrP: a Comparison with Italian

We have seen that BrP verbs which can (or must) denote reciprocity without se in finite clauses can appear in the discontinuous reciprocal construction and allow a reciprocal interpretation with morpho-syntactically singular group NPs, properties that characterize lexical reciprocals in other languages. These data support our hypothesis that verbs which do not require se in reciprocal finite clauses have a lexical reciprocal entry. From these data, it also emerges that the reverse generalization does not fully hold: we identified seven verbs (encontrar ‘to meet’, falar ‘to talk’, corresponder ‘to correspond’, consultar ‘to consult’, unir ‘to merge/combine’, sobrepor ‘to overlap’ and confundir ‘to confuse’) whose behavior with respect to discontinuous reciprocity and group NPs is consistently analogous to lexical reciprocals, although they cannot convey reciprocal interpretations without se in finite clauses. Moreover, these verbs denote concepts that are generally expressed by lexical reciprocals cross-linguistically (Haspelmath, 2007). A possible way to look at this variability in the status of se with verbs that otherwise display a reciprocal behavior is to assume that se is ceasing to be used with lexical reciprocal verbs, and it is bound to disappear with verbs with such an entry. In this view, verbs with an optional se might be in the process of losing it, while for other verbs, the loss of se could be expected in the future; the disappearance of the reflexive se in BrP is noted in many researches (see Bittencourt (2009), a.o.). Furthermore, for the verbs unir ‘to merge/combine’, sobrepor ‘to overlap’ and confundir ‘to confuse’, the obligatory presence of se might be linked to the causative/inchoative alternation that these verbs undergo (see footnote 6).

The non-uniform distribution of se across different lexical reciprocal verbs with the same semantic properties, as well as across different constructions for the same verb (see footnote 5 and 7), validates the existing proposal that se does not directly contribute to the reciprocal interpretation of these verbs, which is equally available regardless of the presence of se; the reciprocal interpretation must therefore originate from the verb root.

This pattern allows a comparison with Italian. As we have seen, Italian verbs with a lexical reciprocal entry equally allow reciprocal readings in causatives (where si is disallowed) and in finite clauses (where si is obligatory). Therefore, si cannot be considered responsible for lexical reciprocal meanings, which are rather due to the lexical meaning of some verbs. This pattern led us to propose that si is a marker of intransitivity, in line with Labelle (2008), and whose obligatory presence is dictated by syntactic requirements of finite clauses (Palmieri et al. 2018). This analysis finds further support in the data from BrP, and it narrows down the difference between Italian and BrP: we propose that the function of si/se is the same in both languages, but they differ in the syntactic requirements for the use of this element in finite clauses. Italian lexical reciprocal verbs can convey a reciprocal interpretation without si in causative constructions, while this element is
always obligatory in finite clauses. On the other hand, in BrP se can be omitted also in finite clauses, in the presence of a lexical reciprocal entry. Thus, in both cases the reciprocal interpretation results from the verb root and does not change regardless of the presence of si/se. The difference between the distribution of si/se in Italian and BrP, therefore, is to be analyzed in terms of syntactic requirements of finite clauses, which are left for further research.

Our second question in this paper concerns the possible correspondence between the meanings that are expressed by lexical reciprocal entries in BrP and Italian. In typological works, it has been proposed that the concepts that are lexicalized as reciprocals constitute a cross-linguistically stable class (Haspelmath, 2007). Therefore, it is natural to expect consistency between lexical reciprocals across these two Romance languages.

In order to draw a comparison, let us first present the Italian lexical reciprocal verbs we identified in Palmieri et al. (2018), reported below in Table 3. These verbs express mutual configurations without si in causatives, while this element is required in finite clauses. The table includes verbs with a transitive binary entry (i.e. where the second reciprocal argument is a direct object), as well as a few verbs which do not have a transitive binary entry, and that were not included in Palmieri et al. (2018): it is the case of the verbs fidanzarsi ‘to get engaged’, scontrarsi ‘to collide’, riconciliarsi ‘to reattach’, scambiarsi qualcosa ‘to exchange something’.

**Italian verbs that require si in finite clauses and have a lexical reciprocal entry**

<table>
<thead>
<tr>
<th>Italian verb</th>
<th>Lexical reciprocal entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>abbracciare</td>
<td>‘to hug’</td>
</tr>
<tr>
<td>lasciare</td>
<td>‘to leave/break up’</td>
</tr>
<tr>
<td>consultare</td>
<td>‘to consult’</td>
</tr>
<tr>
<td>baciare</td>
<td>‘to kiss’</td>
</tr>
<tr>
<td>incontrare</td>
<td>‘to meet’</td>
</tr>
<tr>
<td>coccolare</td>
<td>‘to cuddle’</td>
</tr>
<tr>
<td>salutare</td>
<td>‘to greet’</td>
</tr>
<tr>
<td>sposare</td>
<td>‘to marry’</td>
</tr>
<tr>
<td>frequentare</td>
<td>‘to date’</td>
</tr>
<tr>
<td>inorciare</td>
<td>‘to bump into’</td>
</tr>
<tr>
<td>battere</td>
<td>‘to fight’</td>
</tr>
<tr>
<td>intrecciarsi</td>
<td>‘to intertwine’</td>
</tr>
<tr>
<td>sovrapporsi</td>
<td>‘to overlap’</td>
</tr>
<tr>
<td>confondersi</td>
<td>‘to confuse/blend with’</td>
</tr>
<tr>
<td>alternarsi</td>
<td>‘to alternate’</td>
</tr>
<tr>
<td>unire</td>
<td>‘to merge/combine’</td>
</tr>
<tr>
<td>mescolare</td>
<td>‘to blend’</td>
</tr>
<tr>
<td>fidanzarsi</td>
<td>‘to get engaged’</td>
</tr>
<tr>
<td>scontrarsi</td>
<td>‘to collide’</td>
</tr>
<tr>
<td>riconciliarsi</td>
<td>‘to reattach’</td>
</tr>
<tr>
<td>scambiarsi</td>
<td>‘to exchange something’</td>
</tr>
</tbody>
</table>

Table 3. Italian lexical reciprocal verbs, adapted from Palmieri et al. (2018)

It is worth pointing out that also in Italian there is a class of verbs that give reciprocity without si (18), but whose binary entry requires the preposition con ‘with’ to introduce the second reciprocal argument, similarly to the BrP verbs in Table 1. These verbs have often been overlooked in the literature, but they show properties of lexical reciprocity that we discussed so far: they allow discontinuous reciprocity (19a) and reciprocal reading with group NPs (19b). Table 4 presents a list of these verbs.

(18) Irene e Sara discutono / collaborano
    Irene and Sara discuss.PRES.3P collaborate.PRES.3P
    ‘Irene and Sara discuss/collaborate’

---

8 The verbs fidanzarsi ‘to get engaged’, scontrarsi ‘to collide’, riconciliarsi ‘to reattach’, scambiarsi qualcosa ‘to exchange something’ do not have a transitive binary entry (i), but they necessarily require the presence of si, as well as the preposition con ‘with’ to introduce the reciprocal argument (ii), as in the discontinuous reciprocal construction:

(i) *Irene fidanza Sara (ii) Irene si fidanza con Sara
Irene engage.PRES.1S Sara Irene si engage.PRES.1S with Sara
‘Irene gets engaged with Sara’
A note on lexical reciprocity in Brazilian Portuguese

(19) a. Irene discute / collabora con Sara
    Irene discuss.PRES.3S collaborate.PRES.3S with Sara
    ‘Irene discusses/collaborates with Sara’

b. La squadra discute / collabora
    the team discuss.PRES.3S collaborate.PRES.3S
    ‘The members of the team discuss/collaborate’

<table>
<thead>
<tr>
<th>Italian verbs that do not allow si in finite clauses and have a lexical reciprocal entry</th>
</tr>
</thead>
</table>

Table 4. Italian lexical reciprocal verbs which do not allow si.

At this point, it is possible to observe a correspondence between lexical reciprocals in BrP and Italian. The comparison between each BrP verb and its Italian counterpart is determined by the translation of these verbs into English. Therefore, we do not claim that verbs that are compared here have exactly the same interpretation, but rather that they generate analogous readings and denote approximately the same realm of meanings.

As Table 5 shows, there is no complete overlap, but a remarkable consistency in the verbs that have a lexical entry in the two languages. Table 5 contains a summary of the verbs that have a lexical reciprocal entry only in one language (i) or in both languages (ii).

<table>
<thead>
<tr>
<th>BrP</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) parecer ‘to resemble’, fofocar ‘to gossip’, namorar ‘to be partner with’, encaixar ‘to match’</td>
<td>coccolare ‘to cuddle’, salutare ‘to greet’, legare ‘to become attached to’, battere/combattere ‘to fight’, frequentare ‘to date’</td>
</tr>
</tbody>
</table>
consultar/consultare ‘to consult/confer’, encontrar/incontrare ‘to meet’,
corresponder/corrispondere ‘to correspond’, misturar/mescolarsi ‘to blend’,
alternar/alternarsi ‘to alternate’, unir/unire ‘to merge/combine’, sobrar/sovrapporre ‘to overlap’, confundir/confondere ‘to confuse’

Table 5. Lexical reciprocal verbs in BrP and Italian.

To conclude, we have observed a correspondence both in the way that reciprocal meanings are conveyed in BrP and Italian, as well as in the class of meanings of lexical reciprocal entries. We propose that both in BrP and Italian, the reciprocal interpretation is due to the verb root of lexical reciprocals, and si/se does not semantically contribute to this interpretation. While the comparison with Italian was motivated by the availability of data on lexical reciprocity in this language, a comparison including more Romance languages could lead to relevant generalizations regarding which concepts are lexicalized as reciprocals, as well as how these meanings are conveyed.

4 A few Words on Grammatical Reciprocity

In the previous section, we proposed that in BrP se does not contribute to the reciprocal interpretation of lexical reciprocal verbs, and that these verbs denote reciprocal configurations due to the meaning of the verb root. The primary focus of this paper is lexical reciprocity, but a main remaining question concerns the contribution of se to grammatical reciprocity: if se does not contribute to lexical reciprocity, what is its role when associated to transitive verbs without any lexical entry? On the one hand, it could be possible to expect a syncretism of se, along the lines of Doron & Rappaport Hovav (2009). According to this proposal, se is a marker when associated with lexical reciprocals, but a pronoun with other transitive verbs. However, in this section, we will illustrate that the data from BrP are in line with our proposal in Palmieri et al. (2018) that si/se can never be considered as a pronoun, and should be analyzed as a marker also when it is associated with transitive verbs.

In Italian, si is always obligatory in finite clauses, even when it coappear with the operator a vicenda ‘mutually, in turns’ (20a). This latter element, however, can lead to a reciprocal interpretation in causatives, where si is disallowed (20b). BrP shows a similar pattern in finite clauses: se and the operator um/(a) a/o outro/a ‘one another’ can coappear (21b) or be in complementary distribution, as in (21a) and (21c), therefore also allowing grammatical reciprocal meanings without se (21c).

(20) a. Irene e Sara *(si) puniscono (a vicenda)
   Irene and Sara SI punish.PRES.3P mutually
   ‘Irene and Sara punish each other’

   b. Ho fatto (*si) punire (*si) Irene e Sara a vicenda
   Have.AUX.1S make.PP SI punish.INF SI Irene and Sara mutually
   ‘I caused Irene and Sara to punish each other’
The distribution of *se* in (21) suggests that this element cannot be considered a reciprocal operator such as English *each other*. In fact, the possible co-occurrence of *se* and *um(a) a/o outro/a* ‘one another’ excludes the possibility that they are both operators. Moreover, considering *se* the reciprocal operator and *um(a) a/o outro/a* an intensifier would not explain the cases where the latter conveys reciprocity without *se* (21c). A possible explanation for these data is that si/se is always a marker, even when associated with transitive verbs. In absence of an overt grammatical reciprocal operator, *si/se* can also license covert reciprocity, as in (20a) or (21a). This proposal is compatible with our discussion of lexical reciprocity: unlike Italian, BrP does not require *se* to obligatorily appear in finite clauses, if another source of reciprocity is present (such as a lexical reciprocal entry or, in this case, a grammatical operator).

Also in this respect, further research on the nature of *si/se* and the restrictions on its distribution might be needed to provide further support in favor of this proposal.

## 5 Conclusions

Unlike other Romance languages, BrP does not always require the presence of *se* in order to convey reciprocal meanings in finite clauses. In this paper, we looked at the distribution of *se* and focused on two main questions. On the one hand, we asked whether the BrP verbs that can generate a reciprocal interpretation without *se* have a lexical reciprocal entry. On the other hand, we investigated to what extent the class of lexical reciprocals in BrP overlaps with the one identified in Italian, a Romance language which expresses reciprocity in a more ‘conventional’ way, i.e. where *si* is obligatory in finite clauses.

We proposed that BrP verbs that can give a reciprocal interpretation without *se*, do have a lexical reciprocal entry, and we supported this claim with evidence from different properties that characterize lexical reciprocals in other languages: (i) the discontinuous reciprocal construction and (ii) the availability of reciprocal readings with morpho-syntactically singular group NPs. We have seen that we cannot generalize the obligatory presence of *se* as lack of a lexical reciprocal entry: we encountered some verbs which show semantic properties of lexical reciprocity, although they cannot appear without *se* in finite clauses. Moreover, we have shown that there is a great correspondence between verbs that have a lexical reciprocal entry in BrP and Italian.

From a closer comparison between BrP and Italian, it also emerged that despite the surface differences, these languages do not differ substantially in the way they convey reciprocal meanings. We proposed that the difference relies on the obligatory presence of *si/se* in finite clauses, whose nature is purely syntactic. In none of these languages there seems to be convincing evidence to consider *si/se* responsible for reciprocal meanings, neither lexical nor grammatical. Further research on the cross-linguistic distribution of *si/se* might elucidate the role of this element from a syntactic perspective. Moreover, extending the research to other Romance languages could help establish whether the lexical semantic distinctions found in BrP and Italian can be generalized.
References


1 Introduction

Reciprocal verbs like *fight*, *talk* and *meet* involve three different argument realizations that lead to contrasts as in the following sentences:

(1) a. Sue fought the disease.
    b. ?Sue fought with the disease.
    c. #Sue and the disease fought.

The inanimate object in the transitive sentence (1a) is fully acceptable. By contrast, the reciprocal intransitive form (1c) exhibits a strong selectional violation (‘#’). The acceptability (‘?’) of the *with* construction (1b) is somewhere in between (1a) and (1c). Similar three-way distinctions have also been observed in languages where the class of verbs that exhibit reciprocal alternations is much richer than in English (see Rákosi 2008 on Hungarian, Bar-Asher Siegal 2015 on Hebrew). Rákosi proposes a straightforward approach to contrasts as in (1), postulating different thematic roles for different argument positions. Thus, while the role of *the disease* is standardly ‘Patient’ (or ‘Theme’) in (1a) and ‘Agent’ in (1c), its position in (1b) is assigned a different role. Rákosi refers to this thematic role as ‘Partner’. The contrasts in (1) are then described using the assumption that different thematic roles lead to different verb meanings and introduce different selectional restrictions.

While this is surely a proper description, it must be semantically elaborated. We should like to have a systematic account of the way selectional restrictions as observed in (1) are related to the meaning of parallel sentences without any selectional violation:

(2) a. Sue fought Dan.
    b. Sue fought with Dan.
    c. Sue and Dan fought.

*The work on this paper was funded by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 742204). I’m grateful to Stavroula Alexandropoulou, Elitzur Bar-Asher Siegal, James Hampton, Ed Keenan, Imke Krutwagen, Rick Nouwen, Giada Palmieri, Eva Poortman, Malka Rappaport-Hovav, Anna Szabolcsi and Joost Zwarts for their help. I’m especially indebted to Edit Doron, whose wisdom and friendship I will always miss.*
We may reasonably expect the pattern of (non-)violations in (1) to follow from the way speakers and hearers understand the sentences in (2), which exhibit no violation. Rákosi (2008:p.424) proposes that certain semantic properties like volition, sentience and being the cause of the event are obligatory for the agent but only optional – though, presumably, still preferred – for the partner. If successful, this idea should allow us to analyze the violations in (1b-c) on the basis of the interpretation of the parallel sentences (2b-c). For the analysis we focus on cases where Sue has all relevant Agent properties – she volitionally and sentiently causes the fighting event, whereas Dan takes part in the event but does not have all these properties. In such situations we expect sentence (2c) to be interpreted as false, whereas (2b) may still be true. This kind of asymmetric situations is the basis for the experimental work in Kruitwagen et al. (2017) on the semantics of reciprocal verbs. In Kruitwagen et al.’s work, truth-value judgements of speakers on sentences like (2a-c) in asymmetric situations serve as core pieces of evidence about the semantics of reciprocal alternations. On the basis of their experimental results, Kruitwagen et al. follow Rákosi and characterize the semantic features of verbs according to their salience for different arguments. This salience of semantic features can be expressed on a scale between obligatory (or highly salient), through preferential, to contingent (=fully optional). This idea harmonizes well with Dowty’s (1991) approach to thematic roles as bundles of prototypical properties, as well as with theories of lexical concepts in cognitive psychology (Hampton, 2007, Hampton and Winter, 2017), where vagueness, graded memberships and family resemblance are core theoretical notions. The aim of the present paper is to develop further the conceptual semantics of reciprocal verbs, so to account for contrasts in selectional restrictions as illustrated in (1).

The paper is structured as follows. Section 2 reviews some relevant cross-linguistic data and introduces a generalization about the pattern of selectional restrictions with reciprocal verbs. Section 3 develops ideas from previous work that treat selectional restrictions as presuppositions (McCawley, 1968, Asher, 2011). More specifically, we treat selectional restrictions as following from a general semantic scheme that presupposes the possibility of predication. Section 4 proposes that contrasts in selectional restrictions as in (1) follow as a corollary from the presuppositional treatment of selectional restrictions in Section 3 and a theory of reciprocal alternations, relying on experimental results from Kruitwagen et al. (2017) and the proposal in Winter (2018). It is concluded that a well-developed theory of reciprocal alternations and the lexical concepts that they involve may directly account for the selectional restrictions on different realizations of the verb’s argument. It is hypothesized that a similar conclusion may hold with other types of verb alternations and selectional restrictions.

2 The Reciprocal Selection Generalization

The examples below demonstrate a cross-linguistic pattern with selectional restrictions, of the same type that is illustrated by (1) above.

*English:*

(3) a. Sue talked/gossiped/chatted/made love to/with a tree.
   b. #Sue and the tree talked/gossiped/chatted/made love.

(4) a. ?The drunk *conversed* with the tree.
   b. #The drunk and the tree *conversed.*

---

1For a related idea on preferential symmetry see Gleitman et al. (1996), Partee (2008), remarked on in note 12 below.
(5) a. Dan agreed with the statement.
   b. #Dan and the statement agreed.

(6) a. Sue collided with the tree.
   b. #Sue and the tree collided.

**Dutch** (Imke Kruitwagen and Joost Zwarts, p.c.):

(7) a. Suus vocht tegen de slaap.
   Suus fought against the sleep
   “Suus fought against her sleep”
   b. ?Suus vocht met de slaap.
   Suus fought with the sleep
   “Suus fought with her sleep”
   c. #Suus en de slaap vochten.
   Suus and the sleep fought
   “Suus and her sleep fought”

(8) a. Jan heeft het kussen geknuffeld.
   Jan has the pillow hugged
   “Jan hugged the pillow”
   b. ?Jan heeft met het kussen geknuffeld.
   Jan has with the pillow hugged
   Roughly: “Jan behaved as if he and the pillow hug”
   c. #Jan en het kussen hebben geknuffeld.
   Jan and the pillow have hugged
   “Jan and the pillow hugged”

(9) a. De autochauffeur is tegen/met de boom gebotst.
   the driver is against/with the tree collided
   “The driver hit/had a mutual collision with the tree”
   b. #De autochauffeur en de boom zijn gebotst.
   the driver and the tree are collided
   “The driver and the tree collided”

**Hebrew**:

(10) a. ha-shikor xibek et / nishek et / laxash l- / latef et ha-pesel.
   the-drunk hugged ACC / kissed ACC / whispered to / caressed ACC the-statue
   “the drunk hugged/kissed/whispered to/caressed the statue”
   b. ?ha-shikor hitkabek / hitnashek / hitlaxesh / hitlatef im ha-pesel.
   the-drunk hugged / kissed / whispered / exchanged caresses with the-statue
   Roughly: “the drunk behaved as if he and the statue hug/kiss/whisper/exchange caresses”
   c. #ha-shikor ve-ha-pesel hitxabku / hitnashku / hitlaxshu / hitlatfu.
   the-drunk and-the-statue hugged / kissed / whispered / exchanged caresses
   “the drunk and the statue hugged/kissed/whispered/exchanged caresses”
(11) a. ha-matos hitnagesh b-?im ha-karka.²
the-plane collided in/with the-ground
“the plane hit/had a mutual collision with the ground”
b. #ha-matos ve-ha-karka hitnagshu.
the-plane and-the-ground collided
“the plane and the ground collided”

Hungarian (Rákosi 2008 and Anna Szabolcsi p.c.):

(12) a. János csókol-gat-ta a szobor-t.
Janos kissed-ITER-PAST the statue-ACC
“Janos kissed the statue repeatedly”
Janos kissed-RCP-ITER-PAST the statue-WITH
“Janos engaged repeatedly in a mutual kiss with the statue”
c. #János és a szobor csókol-óz-gat-ott (egymással).
Janos and the statue kissed-RCP-ITER-PAST (with each other)
“Janos and the statue kissed repeatedly (with each other)”

Greek (Dimitriadis 2004 and Stavroula Alexandropoulou p.c.):

(13) a. O Nikos filise to aghalma.
the Nick kissed the statue
“Nick kissed the statue”
b. ?O Nikos filithike me to aghalma.
the Nikos kissed-RCP-SG with the statue
“Nick engaged in a mutual kiss with the statue”
c. #O Nikos kje to aghalma filithikan.
the Nick and the statue kissed-RCP
“Nick and the statue kissed”

From these examples we derive the following generalization:

²Despite empirical claims to the contrary in (Siloni, 2012:p.299), examples with inanimate objects following the Hebrew form hitnagesh im ‘collided with’ are not ruled out by Hebrew the speakers I consulted. For instance:
(i) nahag ha-masa’it asher hitnagesh im ha-gesher shel kvish 4.
   driver the-truck that collided with the-bridge of road 4
   “the truck driver who hit the bridge over Road 4” (http://sharonsharaby.blogspot.com/2017/10/7.html Accessed: 6 July 2019)

Acceptability here is similar to other cases of inanimate objects following Hebrew im ‘with’, as illustrated above and in Bar-Asher Siegal (2015). Similar examples to (i) are found online with inanimate nouns like ‘wall’, ‘ground’, ‘frame’, ‘cubes’, ‘house’ and ‘pole’. Siloni is probably right that the form hitnagesh b- (“hit”, lit. ‘collided in’) is preferred in such cases, which is supported by Google counts, but such preferences do not seem to follow from a grammatical rule.
(14) **The Reciprocal Selection Generalization (RSG):** Let verb\(_1\), verb\(_2\) and verb\(_w\) be three forms of a reciprocal verb:\(^3\) an intransitive-collective form, a binary (transitive/"to"/"against") form and a ‘with’ form, respectively. Observing that the subject SRs of the three forms are the same, we let N be a noun that is acceptable as heading the object of verb\(_2\) form, but not the subject of any of the forms. We have the following acceptability scale:

\[
\text{verb}_2 \ N \ > \ ?\text{verb}_w \ N \ > \ \#\text{N} \ \text{verb}_{1/2/w}
\]

In words: N is more acceptable in objects of the binary form than it is in ‘with’ PPs, which are in turn more acceptable than cases where N appears in the subject position (of any of the forms).

Accounting for the RSG first involves having an account of SRs in general. This is the subject of Section 3. The next step is to derive the SRs of reciprocal verbs from their meanings. This involves accounting for the fact that SR violations that are triggered by the ‘with’ form are milder than those that are triggered by the subject, although still noticeable compared to the object of the binary form. This step will be taken in Section 4.

### 3 Selectional restrictions as presuppositions

To treat the selectional restrictions (SRs) of reciprocal verbs, it is useful to first make some assumptions about the nature of SRs in general. For example, let us consider the following simple case of SR violation:

(15) #Fido drank the meat.

In linguistics there is a fairly long tradition assuming that violations as in (15) directly follow from the verb’s meaning. Thus, understanding what *drink* means – say, to take liquid into the mouth and swallow it – should lead speakers to consider (15) unacceptable, provided that they, reasonably, do not categorize meat as liquid. This line of account also explains why SRs support semantic entailments as in (16):

(16) Fido drank what was left in the bowl ⇒ What was left in the bowl was liquid

In (16), the SRs of *drink* allow us to conclude that the pseudo-cleft *what was left in the bowl* refers to a liquid.

SRs as illustrated (15) and (16) act as parts of the verb’s meaning, which we consider as presuppositions of that meaning (McCawley, 1968, Asher, 2011). One reason for this classification is that SRs show projection behavior that is typical of presuppositions and not of other types of inference. This is observed with SR violations as in the following sentence:

(17) #If Fido drank the meat he got poisoned.

The SR violation in (17) is embedded in the antecedent of the conditional. This violation does not lead here to a conditional statement like “if the meat is liquid and Fido took it into his mouth and swallowed it, he got poisoned”. Rather, the violation gets “projected” as unacceptability of the embedding conditional, with the infelicitous implication about liquid meat. We observe the same projection behavior in setups where SRs are not violated. Consider for instance the felicitous conditional sentence (18), and the two possible conclusions from it in (18a-b):

\(^3\)By referring to the verb as “reciprocal”, we assume that the meanings of the sentences \(A \text{ and } B \text{ verb}_1\), \(A \text{ verb}_2\) \(B \text{ and } B \text{ verb}_2\ A\), and \(A \text{ verb}_2 \text{ with } B\) (\(B \text{ verb}_2 \text{ with } A\)) are roughly the same. Cases of comitative with (*Sue ate with Dan*) and instrumental with (*Sue ate the pizza with a fork*) are thus ignored.
(18) If Fido drank what was left in the bowl he got poisoned.
   a. ⇒ What was left in the bowl was liquid.
   b. ⇐ Fido took what was left in the bowl into his mouth.

The sound entailment in (18a) again demonstrates that the “liquid” SR gets projected as a conclusion of the conditional sentence, this time without leading to any infelicity. By contrast, (18b) demonstrates that the conditional does not entail the “take into mouth” part of the meaning of drink. This is a classical distinction between presuppositions, which “project” out of antecedents of conditionals, and other inferences, which do not (Chierchia and McConnel-Ginet, 1990). Thus, we conclude that SRs are presuppositional, unlike other ingredients of verb meaning.4

Having seen some presuppositional properties of SRs, let us now address the origins for their difference from other ingredients of verb meaning. For example, what principle determines that the “liquid” part of the meaning of drink should be an SR (hence a presupposition), whereas the “take into mouth” ingredient should be part of what is being asserted?5 Rick Nouwen (p.c.) proposes a simple general answer to this question. According to Nouwen’s proposal, SRs manifest a possibility presupposition about predication. Importantly, this proposed presupposition is assumed to be a general formal scheme that concerns all ingredients of a verb’s meaning, not just SRs. For instance, let us assume that the core meaning of drink contains all conceivable semantic ingredients of that verb’s meaning, as formalized in (19) below:

(19) **Core meaning:**

\[ \text{DRINK}_C(e, x, y) = \text{liquid}(y) \land \text{take\_into\_mouth}(e, x, y) \land \text{swallow}(e, x, y) \]

In words, the core meaning of the verb drink requires: x drinks y in an event e if and only if y is liquid, and e is an event where x takes y into her mouth and swallows y.

This does not give us yet any SR. Nouwen’s proposal is that the SRs of a verb are systematically derived from its core meaning by applying a possibility operator. Thus, for the verb drink we get:

(20) **Presuppositional meaning:**

\[ \text{drink}_P(e, x, y) = \Diamond (\exists e'. \text{DRINK}_C(e', x, y)) \]

In words: the presuppositional part of the meaning of drink is a statement about the possibility of having an event where the core meaning holds of the thematic arguments. Based on the core meaning in (19), we conclude:

(21) \[ \text{drink}_P(e, x, y) = \Diamond (\exists e'. \text{liquid}(y) \land \text{take\_into\_mouth}(e', x, y) \land \text{swallow}(e', x, y)) \]

The presupposition in (21) still seems quite far from the SRs that are actually manifested with the verb drink. Specifically, the clauses \text{take\_into\_mouth}(e', x, y) and \text{swallow}(e', x, y) require the possibility of “take into mouth/throat”, which are not clearly manifested as SRs of drink. By contrast, the clause about the possibility that y is liquid seems to require too little: in usages of drink, the object is actually, not just possibly, required to be liquid. To get an idea of the full picture we need to also take into account additional semantic factors:

---

4There are familiar presuppositional ingredients of verb meanings that are not standardly classified as SRs. For instance, the “had smoked” ingredient of aspefical verb phrases like stopped smoking and continued smoking is presuppositional, unlike the “no longer smokes” and “still smokes” ingredients of their respective meanings, which are part of their contribution to assertion.

5This is the triggering problem for presuppositions (Schwarz, 2019), when applied to SRs.
(22) (i) The predicate liquid is a stative predicate, which is presumably constant across the possible situations that are relevant for (21). Thus, something that is required to “possibly” be a liquid may be forced to actually be liquid, since liquidity (like any other stative property) is part of what is intensionally constant about an entity. Thus, an ice cube is viewed as epistemically different from the water that was put in the freezer.

(ii) The eventive predicates take_into_mouth(e, x, y) and swallow(e, x, y) have their own stative SRs: have_mouth(x) and have_throat(x), respectively.

(iii) Once the three stative SRs “liquid”, “mouth” and “throat” are satisfied, we may reasonably assume that the modal statement in (21) is trivially satisfied: any agent x that has a mouth and a throat can possibly take any liquid y into her mouth and swallow it.

It would take us too far afield to derive (22i-iii) from more basic semantic or pragmatic assumptions. However, I take it that they are plausible enough, and for one thing, they are sufficient for our current purposes. Based on these three points, we assume that Nouwen’s proposal can derive the following actual meaning for drink using its core meaning in (19):

\[
\text{drink}(e, x, y) = \\
\text{liquid}(y) \land \text{have_mouth}(x) \land \text{have_throat}(x) : \text{take_into_mouth}(e, x, y) \land \text{swallow}(e, x, y)
\]

With the presuppositional and assertive elements separated by ‘;’, the formula in (23) expresses the proposal that stating that x drinks y in an event e presupposes that y is liquid and that x has a mouth and a throat, and asserts that e is an event where x takes y into her mouth and swallows y.\(^6\)

More generally, we assume that each verb has a core Davidsonian meaning \(\text{VERB}_C\), which maps any event and entity arguments to a truth-value. This core meaning is mapped to an actual meaning \(\text{verb}\) using Nouwen’s principle:

\[
\text{verb}(e, x_1, \ldots, x_n) \\
= \lambda e. \lambda x_1. \ldots \lambda x_n. \Diamond(\exists e'. \text{VERB}_C(e', x_1, \ldots, x_n)) : \text{VERB}_C(e, x_1, \ldots, x_n)
\]

In words: a verb’s actual meaning \(\text{verb}\) applies its core meaning \(\text{VERB}_C\) to the event and thematic arguments, with a possibility presupposition triggered by that core meaning.

Importantly, “core meanings” as tentatively formalized in (19) should not be seen as logical definitions. Rather, they should be seen as meaning postulates, or semantic templates (Rappaport-Hovav and Levin, 1998). The scheme in (19) establishes a connection between the meaning of the verb drink and the meanings liquid, take_into_mouth and swallow. Schemes like that encode assumptions about concepts, whose detailed study is a big enterprise, only partially linguistic (Laurence and Margolis, 1999). For our purposes here, two general properties of concepts are important: their fuzzy boundaries and reliance on non-essential properties. Fuzzy boundaries are illustrated by vague adjectives like red, loud or expensive. Obviously, it is impossible to define once and for all what counts as red, loud or expensive. The same holds for verbs. What are the circumstances under which you would say that the

\(^6\)The formula (19) takes \(\text{liquid}(y)\) to be part of the core meaning of \(\text{drink}\), from which the corresponding presupposition is derived by Nouwen’s proposed principle. Once a stative predicate like that is part of a verb’s meaning, its treatment as a presupposition masks its assertive content. For a similar point about “masked assertive contents” in other cases see Klinedinst (2016), Zehr and Schwarz (2018).
Non-essential properties are another well-known aspect of concepts, at least since the classical works by Wittgenstein (in Philosophy) and Rosch (in Cognitive Psychology). The classes of objects that we categorize as games, fruits or furniture not only have fuzzy boundaries, but are also determined using complex considerations of family resemblance. A well-known example by Wittgenstein is the concept GAME, which involves competition as a common salient property. However, competition is not an essential property of games: many activities that are classified as games are not considered competitive (a Google search yields impressive lists of such games). Thus, we say that “being competitive” is a preferential property of the concept GAME. Other properties of this concept, like “improve skills”, “relieve boredom” or “entertain” are similarly non-essential. In this respect, the meaning of a verb like play are not less multi-dimensional than that of the noun game: to the extent that the verb concept PLAY is related to the noun concept GAME, it must involve non-essential but preferential properties. Such preferential properties are distinguished from a property like “take something into mouth” of the verb concept DRINK, which is prominent enough to be considered as essential for practical purposes.

These remarks come to hint at the possibility that principles of meaning alternation with reciprocal verbs as illustrated in Section 2 may also involve fuzzy boundaries and non-essential properties. The next section argues that this is indeed the case, which has direct implications for the analysis of SRs with reciprocal predicates.

4 Selectional restrictions with reciprocal alternations

In this section we get back to the Reciprocal Selection Generalization of Section 2, with the aim of accounting for it using the general ideas about selectional restrictions in Section 3. To do that, we have to spell out the meaning of intransitive verbs like fight, hug and collide, and analyze their semantic relations with the corresponding binary forms. Much previous work has assumed that such collective verbs require strong reciprocity in the sense of Dalrymple et al. (1998). Thus, for instance, the sentence Sue and Dan hugged is supposed to require that the sentence Sue hugged Dan and Dan hugged Sue holds in one “semantically irreducible” event (Dimitriadis, 2004, 2008, Siloni, 2012). However, recent experiments by Imke Kruitwagen cast doubts on the assumption that reciprocal intransitive verbs generally require this kind of symmetric participation (Kruitwagen et al., 2017, Kruitwagen, 2019). Consider for instance the following Dutch examples:

(25) a. Violet en Mark zijn gebotst.
    “Violet and Mark collided"

b. Mark is tegen Violet gebotst.
    “Mark is against Violet collided"

According to previous accounts, we might expect binary statements as in (25b) to be necessary conditions for the truth of collective sentences like (25a). Kruitwagen (2019) tested this expectation on two groups of 59/53 Dutch speakers, who were requested to make a truth-value judgement on sentence (25a/b) respectively. Each participant was shown a short video film where Violet rides her bicycle and hits Mark’s bicycle, while Mark is standing still. 69% of the participants in the first group accepted sentence (25a) as true in this situation, while only 4%
of the participants in the other group accepted (25b) as true. From such results we conclude that for the majority of speakers, symmetric participation is not necessary for judging sentences like (25a) as being true. Kruitwagen’s experiments show similar results with the verbs *knuffelen* ‘hug’, *vechten (tegen)* ‘fight (against)’ and *fluisteren (tegen)* ‘whisper (to)’. The conclusion is that for many speakers symmetric participation is not obligatory with reciprocal verbs, at least not in certain circumstances.

A weaker logical requirement that is empirically sounder than symmetric participation is what we may call “disjunctive participation”. For two objects to be considered as fighting, hugging or colliding, a necessary requirement is that at least one of them is fighting, hugging or hitting the other. Spelling out a minimal requirement for the core intransitive meaning of *collide*, we get the following disjunctive rule:

\[ \text{Requirement from core meaning of intransitive *collide/botsen:*} \]

\[
\text{COLLIDE}_C(e, x+y) \Rightarrow \text{hit}(e, x, y) \vee \text{hit}(e, y, x)
\]

In (26) we assume that the agent denotes a sum \(x+y\) of two entities (Link, 1983). In words, (26) states that an event \(e\) is a collision with the sum \(x+y\) as its agent only if \(x\) hits \(y\) in \(e\) and/or \(y\) hits \(x\) in \(e\). The kind of disjunctive requirement is weak enough so that, arguably, it describes one part of the meaning of reciprocal verbs that all speakers adhere to in all circumstances (Winter, 2018).

Despite its partiality, the disjunctive rule in (26) is still general enough to account for one of the puzzles we encountered with the SRs of reciprocal verbs. Consider for instance the following examples:

\[
\text{(27) a. Sue collided with the wall.}
\]
\[
\text{b. #Sue and the wall collided.}
\]
\[
\text{c. #The wall collided with Sue.}
\]

While sentence (27a) is fully acceptable, sentences (27b) and (27c) are not. On a first blush, this pattern may seem inconsistent with Kruitwagen’s experimental findings. According to Kruitwagen’s results, we expect many speakers to accept the sentence *Sue and Dan collided* if Sue collided Dan. Why is (27b) nevertheless infelicitous for all speakers independently of whether Sue collided with the wall? Apparently, the fact that Dan could have hit Sue while the wall could not makes all the difference.

The treatment of SRs in Section 3 as possibility presuppositions follows this simple intuition. According to that proposal, the core meaning \(\text{COLLIDE}_C\) in (26) is not yet the actual meaning of intransitive *collide*. To have the actual meaning of verbs, we have to take SRs into account, that is consider the presuppositions that emerge when we apply the possibility operator to the core meaning. In the case of the core meaning (26) of intransitive *collide*, this leads to the following presupposition:

\[
\text{The concept that is here denoted \(\text{hit}\) could also be denoted *collide with* with no substantial difference in intention. However, for purposes of cross-linguistic comparison, that might be misleading because in other languages, usages of expressions like *botsen met* in Dutch (lit. ‘collide with’, meaning roughly: ‘come into a mutual collision with’) usually involve more participation from the indirect object than in English. For this reason I here use the label \(\text{hit}\) for the relevant binary concept. For a more general discussion of the use of ‘with’ in different languages, see below.}
\]

\[
\text{To simplify things, the discussion here ignores collective subjects with more than two members. The analysis could be extended to such cases using a requirement like \(\text{COLLIDE}_C(e, X) \Rightarrow \exists x, y \in X. \neg x = y \wedge \text{hit}(e, x, y)\), where \(X\) is a sum of two or more entities.}
\]
(28) **Presupposition of intransitive collide/botsen:**

\[
\text{collide}_p(e,x+y) = \diamond (\exists e'. \text{COLLIDE}_C(e',x+y)) \\
\Rightarrow \diamond (\exists e'. \text{hit}(e',x,y) \lor \text{hit}(e',y,x))
\]

In words: when intransitive collide applies to a sum \(x+y\), its SRs require the possibility that \(x\) hit \(y\) or \(y\) hit \(x\). Thus, the SR of sentence (27b) entails the following statement:

(29) #It is possible that [Sue hit the wall or the wall hit Sue].

Sentence (29) is unacceptable, reasonably due to the projection of the SR violation from the second disjunct (=the wall hit Sue). Thus, given the SR we derive in (28) for sentence (27b), it is not surprising that (27b) is as unacceptable as (27c) and (29) are. According to our account, the actual meaning of intransitive collide presupposes the possibility that each of the agent’s members hits the other. We conclude that each of the agent’s members is presupposed to be a moveable object, hence the SR violation in (27b).

The examples in (30) below show more disjunctive entailments from core meanings of intransitive reciprocals:

(30) \[
\begin{align*}
\text{FIGHT}_C(e,x+y) & \Rightarrow \text{fight}(e,x,y) \lor \text{fight}(e,y,x) \\
\text{HUG}_C(e,x+y) & \Rightarrow \text{hug}(e,x,y) \lor \text{hug}(e,y,x) \\
\text{KISS}_C(e,x+y) & \Rightarrow \text{kiss}(e,x,y) \lor \text{kiss}(e,y,x) \\
\text{TALK}_C(e,x+y) & \Rightarrow \text{talk}_o(e,x,y) \lor \text{talk}_o(e,y,x)
\end{align*}
\]

In a similar way to the account of the SR violation in (27b), the entailments in (30) allow us to explain SR violations in corresponding sentences with reciprocal verbs in different languages, as exemplified above for English fight and talk (1c,3b), Dutch ‘fight’ and ‘hug’ (7c,8c), Hebrew ‘hug’ and ‘kiss’ (10c), Hungarian ‘kiss’ (12c), and Greek ‘kiss’ (13c).

More generally, our account so far explains why certain nouns that are acceptable as objects of a binary verbal form \(verb_2\) are unacceptable in subjects of the reciprocal form \(verb_1\), violating one of the SRs of the subject position. This is one part of the Reciprocal Selection Generalization (14). The other part concerns cases like the English verb fight, which, in addition to their reciprocal intransitive form, have both a binary (transitive) form and a ‘with’ form. As we saw in Section 2, we often find nouns \(N\) that show SR violations in the post-‘with’ position, but these violations are milder than the SRs that appear when \(N\) is promoted to the subject. To analyze the semantic reasons for this behavior, let us again consider one of Kruitwagen’s experiments, this time with the following Dutch sentence:

(31) Violet heef met Mark gevochten.

Violet has with Mark fought

“Violet fought with Mark”

Kruitwagen’s experiment involved a group of 28 Dutch speakers who were requested to make a truth-value judgement on sentence (31). The participants were shown a short video film where Violet attacks Mark violently, while he responds to her verbally but not physically. 75% of the participants accepted sentence (31) as true in this situation. Based on this result, we hypothesize that in the vechten met ‘fight with’ construction, it is preferred that the post-met participant fights the other participant, but this preference is relatively weak, hence it is ignored by many

\[^9\text{As mentioned in (14), these SRs are the same for all three entries: } verb_1, \verb_2 \text{ and } verb_w.\]
participants. As mentioned in Section 3, the existence of such preferential but non-essential properties is a rather common phenomenon with natural concepts.

Given that the activity of the post-‘with’ participant is only a semantic preference, the SRs that are triggered by this preference are expected to be weaker than the SRs triggered by features that are semantically required.\(^{10}\) This accounts for the weakness of the SR violations that are observed for NPs in the position following the ‘fight with’ construction (7b) as well as for other NPs in the post-‘with’ position (1b,8b,10b,11a,12b,13b). By contrast to this weak preference, we have seen that the collective agent of intransitive entries is strictly required to show disjunctive participation. Thus, the SRs that ensue are manifested as strong requirements of intransitive *vechten* ‘fight’ in Dutch (7c) and other similar intransitive reciprocal entries (1c,8c,10c,11b,12c,13c).

The discussion above analyzes the reasons for the Reciprocal Selection Generalization (14). An additional point that does not directly concern SRs of reciprocal verbs can be observed using the following Dutch sentence:

(32) Violet en Mark hebben gevochten.

“Violet and Mark fought”

Truth-value judgements on sentence (32) were tested by Kruitwagen on 53 Dutch speakers, in the same situation where sentence (31) was tested: a film that shows Violet violently fighting Mark, while Mark only responds verbally. In contrast to the 75% acceptability of (31) in this situation, only 40% of the participants accepted sentence (32) as true in the shown film. The theoretical reason for such contrasts is quite clear. When the subject is a singular NP as in the (31), most speakers require it to be active while only preferring the post-*met* participant to reciprocate. When the subject is a conjoined NP as in (32), symmetric participation is still a preference. However, since both participants are now part of a collective agent, the preference that they are both active is stronger than in (31).

The semantic requirements and preferences of the Dutch verb *vechten* ‘fight’ are formally summarized in Table 1. When *vechten* appears with the preposition *tegen* ‘against’, its meaning is similar to that of the English transitive verb *fight*, which we here denote using the binary predicate \(F\).\(^{11}\) When the preposition *met* ‘with’ is used, it is still required that the relation \(F\) holds between the two entities according to the surface form, but now it is also preferred that \(F\) holds in the opposite direction, where the post-*met* NP gets the agentive role. In the intransitive collective form of *vechten*, we assume that no special direction of the fighting is required between the two agents, hence the weaker disjunctive requirement \(F(x,y) \lor F(y,x)\). It is still preferred that both participants are active, hence the conjunctive preference \(F(x,y) \land F(y,x)\). Furthermore, this preference is stronger than in the case of ‘with’, since both participants are now part of one collective agent.

The pattern in Table 1 is proposed here as a general scheme, holding of all reciprocal verbs that have the three forms: the intransitive-collective entry, the ‘with’ form, and a binary entry (transitive or using a preposition). As a rule, with such verbs the ‘with’ form has a meaning that is preferentially (though not logically) symmetric, whereas the other binary form is not

---

\(^{10}\)In the theory of presuppositions, this is the distinction between “weak” and “strong” presuppositions, which is often described using a distinction between “soft” and “hard” triggers (Abusch, 2010).

\(^{11}\)This is a simplification. The meaning of the Dutch verb *vechten* apparently involves more violent action than that of English *fight*. Therefore, we may expect the acceptance rates of the English parallels of sentences (31) and (32) to be somewhat higher than Dutch in the critical situation.
Verb form | Logical form | Requirement | Preference | Strength of preference
---|---|---|---|---
verb + ‘against’ | $\text{fight}_2(x,y)$ | $F(x,y)$ | – | –
verb + ‘with’ | $\text{fight}_w(x,y)$ | $F(x,y)$ | $F(y,x)$ | weak
intransitive | $\text{fight}_1(x+y)$ | $F(x,y) \lor F(y,x)$ | $F(x,y) \land F(y,x)$ | strong

Table 1: semantic requirements and preferences for the Dutch verb vechten ‘fight’

symmetric, even not as a preference. However, not all reciprocal verbs in English have the three forms that fight has. Reciprocal verbs like collaborate, converse and agree do not have a relevant binary entry on top of their with alternate. Another, smaller, class of English reciprocal verbs includes hug, kiss and divorce, which have a transitive form (or another binary form) but no with form. This distinction between verbs is subject to substantial cross-linguistic variation. Reciprocal verbs that have similar meanings in different languages may show differences in the availability of the binary form or the ‘with’ form. Table 2 shows some similarities and differences in this domain between English, Dutch and Hebrew. With the verbs ‘fight’ and ‘talk’ the three languages have both ‘with’ and another binary form. As suggested above, the ‘with’ form in the three languages is more symmetric than the other binary form, and shows weak agentive SRs on the post-‘with’ participant. With the verb ‘collide’, English’s main binary form has the preposition with. Interestingly, English shows much flexibility in the range of nouns appearing in the post-with position, including nouns like wall, bridge and tree that refer to immovable objects. Dutch and Hebrew ‘collide’ allows both ‘with’ and another preposition. Accordingly, these two languages are more restrictive than English concerning the interpretation of the ‘with’ form and the nouns that are allowed to follow it (on Hebrew, see note 2). The verbal concept make love is an opposite case: English allows both with and to, but prefers to for exceptional situations where inanimate objects like trees are involved. Dutch and Hebrew have no colloquial form parallel to make love to, and use ‘with’ for events where people make love to an inanimate object. Another pattern appears with the English verbs kiss and hug, which have no with form, as opposed to the corresponding verbs in Dutch and Hebrew, as well as Hungarian (12) and Greek (13). Accordingly, English has no binary form that expresses preferential symmetry.

Two examples that were mentioned above involve the verbs collide and make love, where the SRs of English with constructions seem to be substantially different than those of Dutch and Hebrew. The tendency that these two cases illustrate is speculated to be rather general:

---

12In an attempt to analyze the factors that determine the class of reciprocal verbs (a question that is not studied here), Gleitman et al. (1996) propose that all binary entries that alternate with reciprocal intransitive forms are symmetric, at least as a preference. While this is apparently the first acknowledgement of the centrality of preferences in the semantics of reciprocity, it is doubtful that the transitive entries of reciprocal verbs like kiss or hug prefer symmetric situations more than transitive entries like resemble or be near, which have no intransitive reciprocal usages in English. Furthermore, Gleitman et al.’s proposal leaves noticeable cross-linguistic gaps: unlike their English parallels, the Dutch verb omhelzen ‘embrace’ has no intransitive entry, whereas the Greek verb miázun ‘resemble’ is used as an intransitive verb to describe a group of similar objects. As argued by Haspelmath (2007), the factors that affect the determination of lexically reciprocal verbs in different languages require much further research.
Selectional Restrictions and Reciprocal Alternations

Table 2: some different expressions of reciprocal verbs in English, Dutch and Hebrew

<table>
<thead>
<tr>
<th>Verb</th>
<th>English</th>
<th>Dutch</th>
<th>Hebrew</th>
<th>Translations (Dutch, Hebrew)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGHT</td>
<td>with/ACC</td>
<td>‘with’/’against’</td>
<td>‘with’/’in’</td>
<td>vechten, nilxam</td>
</tr>
<tr>
<td>TALK</td>
<td>with/to</td>
<td>‘with’/’against’</td>
<td>‘with’/’to’</td>
<td>praten, diber</td>
</tr>
<tr>
<td>COLLIDE</td>
<td>with/-</td>
<td>‘with’/’against’</td>
<td>‘with’/’in’</td>
<td>botsen, hitnagesh</td>
</tr>
<tr>
<td>MAKE LOVE</td>
<td>with/to</td>
<td>‘with’/-</td>
<td>‘with’/-</td>
<td>liefde bedrijven, asa ahava</td>
</tr>
<tr>
<td>GET MARRIED</td>
<td>-/to</td>
<td>‘with’/-</td>
<td>‘with’/-</td>
<td>trouwen, hitxaten</td>
</tr>
<tr>
<td>KISS</td>
<td>-/ACC</td>
<td>‘with’/ACC</td>
<td>‘with’/ACC</td>
<td>kussen/zoenen, nishek/hitnashek</td>
</tr>
<tr>
<td>HUG</td>
<td>-/ACC</td>
<td>‘with’/ACC</td>
<td>‘with’/ACC</td>
<td>knuffelen, xibek/hitxabek</td>
</tr>
</tbody>
</table>

(33) **Conjecture:**

a. If a reciprocal verb allows ‘with’ as well as another binary entry, then the ‘with’ form prefers symmetry – i.e. agentive properties for the post-‘with’ NP. The other binary form shows no preference for symmetry.

   This is the case with English *make love with* and Dutch (Hebrew) ‘collide with/against(in)’.

b. Conversely: a reciprocal verb that only has a binary ‘with’ form may also use it in non-symmetric situations, where the post-‘with’ NP lacks agentive properties.

   This is the case with English *collide with* and Dutch/Hebrew ‘make love with’.

Further examination of this conjecture requires more research.

5 Conclusions

This paper has addressed a puzzle about selectional restrictions (SRs) with reciprocal verbs. It has been observed when the subject of the reciprocal intransitive entry collectively refers to a sum, its agentive SRs strongly apply to each of the sum’s members. Further, when a ‘with’ form coexists with another binary form, agentive SRs of the subject also apply to the post-‘with’ argument of the verb, though in a weaker manner. These observations, which were summarized by the Reciprocal Selection Generalization, raise two questions. First, given that symmetric participation is not required for all intransitive entries of reciprocal verbs, it is surprising that SRs symmetrically holds for the members of the agent’s collective denotation. To account for this behavior, we argued that once SRs are treated as presuppositions of the lexical meanings, their uniform distribution to the collective agent’s members is also expected by the disjunctive semantics that lies at the core of reciprocal verb meanings. Second, we aimed to explain the fact that the agentive SRs triggered by the post-‘with’ argument are relatively weak. This fact as well was explained by the analysis of SRs as projected from lexical meanings. It was proposed that non-essential preferential ingredients of a verb meaning trigger SRs that are weaker than those that are triggered by the verb meaning’s essential elements. Finally, it was conjectured that the strength of agentive properties of reciprocal ‘with’ correlates with the existence of an alternative binary form. It is expected that, to the extent that these findings and theoretical ideas are further substantiated for reciprocal verb alternations they may also be relevant for analyzing semantic facts about other cases of verb alternations.
References


