D. Zelenskii
Lomonosov Moscow State University,
Moscow, 119991, Russian Federation

Universal DP and Bošković phases

This paper is dedicated to creating a restrictive account for syntactic phases under some crucial assumptions of the last years such as Reverse Agree and Universal DP. To achieve that, Bošković’s position on Phase Impenetrability Constraint, which entails that it does not affect Agree, is checked against Russian data reinterpreted under Universal DP.

Key words: phases, syntax, Russian language, Minimalism

Acknowledgments. As for acknowledgments, I am grateful (in chronological order of the last of the influences relevant to this paper) to Ekaterina Lyutikova for teaching syntax, bringing the debate on Universal DP to my attention and co-organizing FARL3, Susanne Wurmbrand for Reverse Agree and overall better understanding of the Minimalist machinery, Sergei Tatevosov for fruitful discussions and co-organizing FARL3, Pavel Iosad for bringing some new papers to my attention and my parents for their neverending support, both emotional and financial, despite having little connection to the world of linguistics.

Д.М. Зеленский

Московский государственный университет имени М.В. Ломоносова, 119991 г. Москва, Российская Федерация

Универсальная группа артикля и фазы Бошковича

Статья посвящена созданию ограничивающего анализа синтаксических фаз при принятии некоторых допущений последних лет, в частности, универсальной группы артикля и обратного согласования. Чтобы достичь этого, позиция Бошковича относительно ограничения непроницаемости фазы, предполагающая, что оно не влияет на согласование, проверяется на данных русского языка, реинтерпретированных с учетом универсальности группы артикля.

Ключевые слова: фазы, синтаксис, русский язык, минимализм

Благодарности. Я благодарю (в порядке времени влияния на статью, от наиболее раннего к наиболее позднему) Екатерину Анатольевну Лютикову за обучение синтаксису, привлечение моего внимания к дискуссии об универсальности DP и соорганизацию конференции FARL3, Сюзанну Вурмбранд за создание механизма обратного согласования и обеспечение лучшего понимания архитектуры минимализма, Сергея Георгиевича Татевосова за плодотворные обсуждения и соорганизацию конференции FARL3, Павла Иосада за привлечение моего внимания к некоторым ранее мне неизвестным работам. Благодарю также моих родителей за их постоянную поддержку – как эмоциональную, так и финансовую – несмотря на отсутствие связи с миром лингвистики.


1. Introduction1

At least since [Chomsky, 2000] phases have been one of the main instruments of a Minimalist linguist in deriving syntactic constraints. They were meant to replace pre-Minimalist barriers of [Chomsky, 1986] and

---

1 While being based on a talk at Third Workshop on Formal Approaches to Russian Linguistics (FARL3), it expands somewhat beyond the talk.
at the same time provide a chunking device, so that derivation happens phase by phase and the computational cost is thus reduced [Chomsky, 2001].

Naturally, nearly all of their properties have been debated. For instance, there has been controversy on whether there is a closed set of phase-inducing heads (such as C, v/Voice,\(^2\) D and P) or phasehood is defined dynamically, as a derivation proceeds; whether phases for PF and LF are the same or different; whether a phase head can have multiple specifiers (see also footnotes); what are the effects of phases (and, in particular, how do they interact with Move (a.k.a. Internal Merge) and Agree; see below); whether they are obligatory or optional (e.g., [D’Alessandro, Scheer, 2015]). Many problems of the phases are discussed in [Boeckx, Grohmann, 2007], though a ban on multiple specifiers hopefully dissolves their argument against phases limiting locality.

The interaction of phases and operations has often been tied to the so-called Phase Impenetrability Constraint (PIC), originally formulated as follows: “In phase \(\alpha\) with head \(H\), the domain of \(H\) is not accessible to operations outside \(\alpha\), only \(H\) and its edge are accessible to such operations” [Chomsky, 2000, p. 108]. [Chomsky, 2001], beyond relaxing it (see below), adds a quasi-explanation: when a phase is completed its complement is transferred to the interfaces (i.e. spell-out), and its internal structure is thus made opaque to the narrow syntax.\(^3\)

Minimalism postulates two\(^4\) main operations: Merge and Agree. Merge comes in several flavors: at least External and Internal Merge are to be distinguished.\(^5\) External Merge obviously cannot target something

---

\(^2\) Chomsky uses \(v\) for the head that does all of the following: projects or does not project external argument (agent), provides or does not provide accusative case for internal argument and serves as a phase for object extraction. While the link between the first two properties has been known for a long time (as “Burzio’s generalization” [Burzio, 1986]) there is no obvious reasons to believe that the last property belongs to the same head. Moreover, the belief leads to an undesirable conclusion that a phase head can have multiple specifiers, which more or less defeats the purpose of replacing barriers with phases. Therefore, it is natural to assume that this is achieved by two different heads – the phase head is \(v\) ([Marantz, 1995]’s verbalizer, same as [Pollock, 1989]’s AgrO) whereas case and external argument are provided by Voice of [Kratzer, 1996]. This also dissolves the long-standing debate on why unaccusatives exhibit phase properties – they do have a verbalizer. Other heads ascribing part-of-speech are presumably phases too, at least when they take root phrases as their complements (it would seem somewhat unpleasant to claim that governmental has two phases, one for ment and one for al).

\(^3\) It obviously cannot be maintained that the structure is fully lost: at least LF seems to have access to it.

\(^4\) Though see [Zelenskii, 2020] for a justification for the third operation of Percolate. Percolate, however, is subject to its own locality limitations, which are much more strict than PIC, and thus irrelevant.

\(^5\) [Citko, 2005] argues at length that Parallel Merge is also needed; it is, however, arguably analogous to External Merge in what concerns PIC and is thus omitted in further discussion.
else than the topmost node and is thus not subject to PIC. Internal Merge, however, also known as Move, involves Merging the topmost node with one of its subconstituents and is therefore subject to PIC – one cannot Move out of a phase without a stop at the phase’s edge. If multiple specifiers are disallowed, this provides an explanation for island effects known since [Ross, 1967].

A more arguable question is whether PIC also holds of Agree. There are at least four answers found in the literature:

1. Yes, unconditionally (original answer of [Chomsky, 2000]);
2. Yes, but only after the next phase is built (so-called weak, or relaxed PIC, originating from [Chomsky, 2001, p. 14], restated in [Richards, 2010]; not to be confused with the notion of weak phases, which are not subject to PIC at all);
3. It depends ([D’Alessandro, Scheer, 2015]);
4. No, unconditionally, all locality limitations on Agree are due to intervention ([Bošković, 2003]).

The last option is the one I will set to explore in this paper – from a very particular point of view. Rather than comparing it to the three other options directly, I shall check whether it is compatible with Universal DP hypothesis (see section 2), which was not accepted by Bošković.

Reverse Agree [Wurmbrand, 2014] (i.e. the one with downward valuation and not linked to interpretability) will be assumed, if only for the simplicity of derivations (though there are conceptual arguments for it as well); note that this is not the one Bošković assumed back in early 2000-ies.

2. The Universal DP debate

Since [Abney, 1987] articles of languages like English have been largely supposed to be a functional head D whose complement is the remaining noun phrase. Soon many properties (including phasehood) were linked to the D projection, and a question arose whether the properties are there in “articless” languages and, if they are, whether they are truly articless (in other words, whether DP is universal).

---

6 Or, to be precise, two topmost nodes of the merged structures, where each is either extracted from the lexicon or built in a separate workspace. The latter situation is found, e.g., for external argument, which is (typically) a DP built in a separate workspace and then attached as a specifier to the VoiceP built in another (“main”) workspace.

7 Otherwise, as it is correctly noted in [Ceplova, 2001] (cit. by [Boeckx, Grohmann, 2007, p. 213–214]), nothing should be an island (without additional undesirable stipulations).

8 Modulo, of course, Saxon genitives, quantifiers surfacing to the left of an article (such as all the men) and other SpecDP elements.
Bošković has been one of the most ardent proponents of the idea that languages lacking overt articles (such as Russian or Serbo-Croatian) indeed lack them (see [Bošković, 2005, 2008]), and the debate shall be reviewed below. I do not intend to claim that I shall mention all the relevant works; however, I maintain that I discuss enough to demonstrate both the situation and the solution.

2.1. Bošković’s arguments against Universal DP

Bošković makes a set of empirical claims (and their theoretical justifications) which are so peculiar that their extended direct quotation is justified:

1. Only languages without articles may allow left-branch extraction;
2. Only languages without articles may allow adjunct extraction out of noun phrases;
3. Only languages without articles may allow scrambling;¹⁰
4. Negation Raising is disallowed in languages without articles (by NPI-test, as interpretation in embedded clause is clearly available);
5. Multiple wh-fronting languages without articles don’t show superiority effects in the fronting;
6. Only languages with articles may allow clitic doubling;
7. Languages without articles don’t allow transitive nominals with two genitives;
8. Only languages with articles allow the majority superlative reading (i.e. reading Most people drink beer as equivalent to There are more people who drink beer than those who do not);
9. Head-internal relatives are island-sensitive in languages without, but not in those with articles (with interpretative differences as well);
10. Polysynthetic languages do not have articles.

Of these, 8 is clearly false¹² (Russian allows the reading for Bol’šinstvo lyudej p’yot/p’yut pivo ‘Most people drink beer’ given lack of stress on pivo ‘beer’; it is quite probable that the languages cited by Bošković were simply not checked for different prosodic patterns), and 10 is ill-defined as it is unclear what is meant by polysynthesis (compare the diversity in Oxford handbook [Fortescue et al., 2017]) – under quite a few definitions French counts as one, but this is obviously not what Bošković has in mind. As for 7,

---

¹⁰ Japanese-style, not German-style – allowing, in particular, scrambling out of dependent finite clauses.
¹¹ In the paper, inexplicably, Negative Raising.
¹² Note that this is the one empirical claim that Bošković does not try to give a theoretical explanation for, so a good devil’s advocate could say that’s why.
Bošković makes an unwelcome stipulation that D feature is involved in movement to SpecCP (which crumbles in too many ways to mention; Cable’s Q [Cable, 2007] being an intervener is enough) to link his otherwise solid analysis [Bošković, 1999] to the articleness.

Additionally, he claims [Bošković, 2008, p. 108–109] that adjectives are also functional heads in languages with articles and adjuncts in languages without, based on “disrupting case assignment” to pronouns in English and Macedonian but not Serbo-Croatian (cf. The real him / *he is dead and Russian Nastoyaschij on myortv ‘idem’) and that determiner-like objects in articless Slavic languages are adjectives (based, in particular, on their morphological behavior, ability to stack up in free order, appearing in predicate positions and inability to be modified).

2.2. Arguments for Universal DP

As the claim of DP-less languages would have huge implications for syntactic theory (such as a direct influence on syntax by morphology, contrary to the standard inverted Y-model both of [Chomsky, 2000] and of previous works), there have always been proponents of the alternative view, namely, that “articless” languages feature a D head which is phonologically zero (or, perhaps, several such heads).

2.2.1. Pereltsvaig’s counter-arguments

Of course, Bošković’s arguments did not go unnoticed. In particular, [Pereltsvaig, 2007] took the pain to thoroughly discuss a lot of them.

The claim that there are no functional heads corresponding to adjectives makes two incorrect predictions. Firstly, the order of adjectives modifying the same noun is expected [Bošković, 2005, p. 6] to be freer (i.e. not subjected to a functional hierarchy) in articless languages, whereas an experimental comparison of Russian and English speakers did not show the effect (and the explicit claim of allowing both of the possible orders actually happened three times as rare for Russian speakers as for English ones). Secondly, the distinction of light (head) vs. heavy (phrasal) adjectival modifiers is predicted to be inexistent, which is, again, false: light adjectives intervene in head movement (namely, approximative inversion) whereas heavy do not.

Likewise, the claim that demonstrative pronouns and prenominal possessives are adjectives does not go through. Morphological similarity is incomplete and, even were it complete, entails nothing about syntactic category. Free ordering does not hold any more than for adjectives. Predicative syntax involves a null noun, as agreement with polite singular vy ‘you’ shows (Vy moya/*moi ‘you are mine’), and, even if it did not, this would no more mean that these words are adjectives than ability of his to appear in a predicative
position (like This book is his) entails that his is adjective.\textsuperscript{13} Modifiability turns out to be an argument against unifying with adjectives, as adverbs can modify adjectives but not these elements. Moreover, possessives actually introduce a referent (which can also bear theta-role of the modified noun), unlike adjectives (even when derived from the same noun).

Left branch extraction is shown to work differently from what Bošković claims, targeting only the adjective itself not necessarily its phrase and not being blocked by other adjectives (but still being blocked by the elements discussed in previous paragraph). Extracted elements need not even constitute a constituent.

The two genitives argument does not go through because the two genitives in languages with overt articles may have different properties (and in English – even different marking: John’s picture of Mary), and, moreover, combination of two low genitives is actually permissible in Russian: konspekt lekcii brata ‘brother’s notes of a/the lecture’. (The situation with (high) possessives is more difficult, see Ch. 2 in [Lyutikova, 2017].)

Scrambling counter-argument in its original form was subverted by the “Japanese-style not German-style” codicil in [Bošković, 2008]. See next section for another argument for Universal DP based on scrambling.

\textbf{2.2.2. Small nominals}

Universal DP hypothesis makes (upon most realizations) a powerful prediction: it is possible to have phonologically identical noun phrases differing in whether they have a DP layer – i.e. whether they are full DPs or so-called small nominals. This is a so-called distributive argument [Lyutikova, 2017, p. 42].

The prediction is borne out, as described at length in [Pereltsvaig, 2006; Lyutikova, 2017] (among others). Here are some examples of their arguments:

Noun phrases used as predicates do not have a DP-layer and thus cannot have demonstrative pronouns in them (although this fact can, in principle, be explained semantically). Vocative positions do not allow DP, either.

Russian quantitative phrases can differ in distribution and agreement depending on presence/absence of DP-layer (1), and the semantics also differs:

\begin{equation}
(1) \text{a. } \left[[\text{Pyat’ xirurg-ov}]_{QP}\right]_{DP} \text{ operirova-l-i Dzhejms-a Bond-a.}
\end{equation}

\begin{itemize}
\item five surgeon-GEN.PL operate-PST-PL James-ACC Bond-ACC
\end{itemize}

‘Five surgeons performed an operation on James Bond’.

\textsuperscript{13} Pereltsvaig uses mine which is segmentally distinct from my; the counter-argument still goes through.
five surgeon-GEN.PL operate-PST-N James-ACC Bond-ACC
‘Five surgeons performed an operation on James Bond’.

c. [Te [pyat’ xirurg-ov]QP]DP operirova-l-i/*-o Dzhejms-a
those five surgeon-GEN.PL operate-PST-PL/-N James-ACC
Bond-a.
Bond-ACC
‘Those five surgeons performed an operation on James Bond’.

DPs require structural case, which is not universally necessary for small
nominals; this also leads DPs (and those small nominals that do, for language-
particular reasons, have structural case, see [Lyutikova, 2014]) to be subject
to scrambling whereas small nominals that lack structural case cannot
undergo it.

DPs are islands, whereas small nominals are not, and the minimally
contrasting pairs can be found in Russian complex predicates such as prinyat’
reshenie ‘make a decision’ whose second part may form NP or DP, with
different possibilities of extraction:

(2) …kotoryj byi-l-o prinya-t-o (*moy-o) reshenie
which be-PST-N make-PST.PASS-N my-N decision
prevrati-t’ v=ehkspluatiru-em-oe pomeschenie
turn.into-INF in=exploit-able-N.ACC place.ACC
‘…which was decided to be turned into an exploitable place’.

Sort inversion is N-to-A-to-D-movement, thus unavailable in contexts
requiring small nominals such as vocatives: *O filin obyiknovennyij!
‘Oh, Bubo bubo!’

2.2.3. Other arguments

Before drawing up a conclusion I shall add two more arguments from
[Lyutikova, 2017] to really “bring the argument home”.

DPs can embed a clausal constituent as well as NP, and the resulting
constituent is an island, as Ossetic is claimed to show [Lyutikova, 2017,
p. 88–117]. This is difficult to model without a DP (and, moreover, a phase DP).

Appositive relative clauses are to be adjoined above DP, whereas restrictive
relative clauses are to be adjoined below; moreover, their relative pronoun
is also different, and the differences are inexplicable without intermediate
projections.
2.3. Conclusion on Universal DP

Although there were many more papers on both sides, it seems that Universal DP proponents make stronger (and more interesting) predictions that tend to be correct, whereas its opponents (and Bošković in particular) rely on some tendencies which are frequently underanalyzed, incorrect or irrelevant – and in a disturbing number of cases some combination of these properties.

It can thus be stated with a decent degree of certainty that Universal DP hypothesis is correct, despite the fact that the burden of proof has been put\(^{14}\) on its proponents.

3. Compatibility

Now, having more or less established that Universal DP is right, I can turn to the question whether irrelevance of PIC for Agree stated by [Bošković, 2003] can be maintained (to the extent it was originally true\(^{15}\)) for “articleless” languages as well given that Universal DP hypothesis rejects the idea of major structural differences between them and languages with overt articles. Russian will be used as a prototypical example.

The type of situations in question is exemplified (3a) below:

(3) a. On pokaza-l [molodozhyon-am]\textsubscript{DP} [podark-i drug=drug-a]\textsubscript{DP}
   he show-pst newlywed-pl.dat gift-pl.acc each=other-gen
   ‘He has shown newlyweds each other’s presents.’

b. On pokaza-l [molodozhyon-am]\textsubscript{DP} drug=drug-a.\(^{16}\)
   he show-pst newlywed-pl.dat each=other-acc
   ‘He has shown newlyweds each other.’

c. ?… [te podark-i drug=drug-a]\textsubscript{DP}
   … those gift-pl.acc each=other-acc
   ‘He has shown newlyweds those presents of each other’s.’

---

\(^{14}\) If the basic ideas of universal and restrictive linguistic capacity are accepted this is actually unexpected. Bošković, moreover, seems to understand this: he notes that “[i]t is often assumed TNP should be treated in the same way in articleless languages and English for the sake of uniformity. This argument fails on empirical grounds: it is simply a fact that there are radical differences between the two–there is no uniformity here” [Bošković, 2008, p. 106]. Thus, he can be understood to agree that the uniform treatment is the null hypothesis and the burden of proof that empirical evidence points otherwise is on him.

\(^{15}\) So it is actually a question of whether Bošković’s arguments on the issue are transferrable, not whether they entail what Bošković wants them to entail. If he is wrong he is, of course, wrong in all languages.

\(^{16}\) The homonymy of accusative and genitive, common to all Russian pronouns (barring chto ‘what’ and its derivatives as well as numerical pronouns such as skol’ko ‘how much’ and gender-agreeing ones) is irrelevant.
As discussed above, the genitive drug=druga ‘each other’ in (3a) is clearly not in SpecDP – at least not overtly. At the same time, the noun phrase containing them is clearly referential – and thus a DP, as stated above. DP’s are (more or less universally) considered to be phases.

The reciprocal pronoun drug=druga in (3a) is clearly bound by molodozhyonam ‘newlyweds.DAT’, given that on ‘he’, being singular (and, on the most natural reading, countable), is an infelicitous antecedent for a reciprocal pronoun. That c-command and similar conditions on binding are established can be seen in (3b), where the same pronoun is bound by the same noun phrase in a minimally differing context without a phase over it.

Binding relations are now generally taken to be instantiations of Agree [Landau, 2000]. Simply put, the binder transfers its φ-features to the bindee (downwards, hence Reverse Agree; standard Chomskian Agree requires additional stipulations). Therefore, as long as we can trust the overt low position of drug=druga to be the highest, (3a) is evidence of Agree into a phase (namely, a DP) in an articleless language. Moreover, as podarki ‘presents’ are to have a nominalizer head over its root, two phases are there 17 between the binder and the bindee (see the second footnote), so the relaxed PIC does not work either.

One could claim that, to the extent that (3a) is grammatical, there is a covert movement to SpecDP (and, indeed, its overall not so great acceptability further declines if SpecDP is occupied, cf. (3c)). Bošković, however, rejects covert movement as a theoretical instrument overall, claiming that having both Agree and covert movement is superfluous (contra [Pesetsky, 2000], as he notes himself). Therefore, to Bošković, (3a) is an argument in favor of his view even under the idea Universal DP.

4. Conclusion

This paper has checked whether arguments of [Bošković, 2003] are compatible with the Universal DP hypothesis and found that they are, indeed, compatible.

This, importantly, does not entail that they are necessarily insurmountable. The debate on the locality of Agree is huge, and findings are somewhat inconsistent (especially when one seeks to unify all the situations which are likely to be instantiations of Agree, from assignment of (structural) case to predicative agreement to binding).

17 Though see [Erschler, 2018] for an opposite view.
As a side result, disallowing multiple specifiers at a phase edge\textsuperscript{18} were justified.

References


\textsuperscript{18} They are probably disallowed overall, though this is much harder to justify since there is no PIC for other situations to render the complement unaccessible and since no island effects are expected there (nor, thankfully, found). [Kayne, 1994] has an explanation for having a single specifier, but it entails accepting a version of linearization that comes off as too strong, leading, among other things, to postulating word order heads having no semantic content, thus defying the principle of full interpretability [Chomsky, 1995, p. 27]. One can claim that PF-interpretability works as well as LF-interpretability, but that would entail that features with no interpretable instances could exist for PF reasons, thus by and large destroying the principle’s main prediction.

Статья поступила в редакцию 10.11.2019, принята к публикации 15.12.2019
The article was received on 10.11.2019, accepted for publication 15.12.2019

Об авторе / About the author

Зеленский Дмитрий Максимович – магистрант кафедры теоретической и прикладной лингвистики филологического факультета, Московский государственный университет имени М.В. Ломоносова

Dmitrii M. Zelenskii – MA student at the Department of Theoretical and Applied Linguistics of the Faculty of Philology, Lomonosov Moscow State University

ORCID: http://orcid.org/0000-0002-1979-0482
E-mail: dz-zd@mail.ru