

КРИТИКА И БИБЛИОГРАФИЯ

ОБЗОРЫ

© 2008 г. П.В. ИОСАД

ТЕОРИЯ ОПТИМАЛЬНОСТИ: ОБЗОР ОСНОВНОЙ ЛИТЕРАТУРЫ

В настоящей работе предлагается обзор основных направлений в рамках теории оптимальности (далее ТО) – фонологической модели, получившей большую популярность в последние 15 лет и нашедшей свое применение не только в фонологии. Показана связь проблематики ТО с вопросами, обсуждавшимися в рамках фонологических теорий 1970–1980-х годов, а также новые направления развития ТО. В частности, в работе указаны проблемы, которые в рамках ТО были впервые поставлены или стали решаться новыми способами, такие как проблема усвоения грамматики из доступных данных или роль обобщений над поверхностными формами. Кратко обсуждаются задачи, к решению которых ТО до сих пор успешно не применена, в особенности проблемы «непрозрачности».

Теория оптимальности (англ. Optimality Theory) – направление, в последнее десятилетие приобретшее большую популярность в западной лингвистике, но в России, к сожалению, практически неизвестное (ср., однако, [Князев 1999]). Между тем именно в рамках теории оптимальности сейчас ставятся и решаются многие проблемы, для которых принятые ранее теории не могли найти удовлетворительного решения – а иногда и вовсе их не ставили всерьез.

Цель предлагаемого обзора – дать читателю представление об основных положениях этой теории, ее сильных (и некоторых слабых) сторонах и направлении ее развития. Он ни в коем случае не претендует на какую-либо полноту (тем более что в некоторых областях полный обзор проблематики ТО совпал бы с обзором всей литературы вопроса за последние десять-пятнадцать лет), но можно надеяться, что он поможет читателю лучше ориентироваться в некоторых активно обсуждаемых в настоящее время проблемах лингвистической теории.

Первая оговорка, которую необходимо сделать, – ТО, вообще говоря, не является собственно фонологической или даже собственно лингвистической «теорией» в том смысле, в какой теориями являются, к примеру, пражская фонология или теория управления и связывания. Теория оптимальности – это скорее некий достаточно общий алгоритм принятия решений, имеющий широкое лингвистическое применение. При этом, будучи алгоритмом, то есть объектом математической теории, он имеет и нетривиальные математические свойства, что позволяет искать новые подходы ко многим давним проблемам (ср. [Broekhuis, Vogel 2006], где обсуждается связь ТО и теории минимализма, а также многочисленные работы по компьютерным приложениям алгоритмов ТО). Тем не менее в дальнейшем изложении мы сосредоточимся именно на лингвистической стороне проблемы, памятуя, однако, о необходимости в некоторых случаях различать ТО как алгоритм и «специальную ТО», то есть его лингвистическое приложение.

Вторая важная оговорка относится не только к теории оптимальности, но и вообще к путям развития фонологии на Западе (да и, вероятно, не только) в последние десятилетия. Развитие фонологии, как, пожалуй, никакой другой лингвистической дисциплины, направлялось в первую очередь работой над весьма конкретными классами задач. По меткому выражению Р. Лэсса, сделанному в его добротном, хотя и несколько устаревшем учебнике [Lass 1984], «phonology is essentially a problem-driven discipline». Так, можно сказать, что в некоторых направлениях структуралистской фонологии такими «направляющими» темами были устройство фонемного инвентаря (ср. хотя бы работы А. Мартине [Martinet 1956]) и природа процесса нейтрализации, в связи с которой можно вспомнить не только дискуссию Московской и Ленинградской фонологических школ, но и работы поздних американских дескриптивистов. Такие «горячие темы» появлялись и позже: после появления основополагающей работы [Chomsky, Halle 1968] на первый план вышел вопрос о допустимой степени абстрактности фонологических представлений, в дискуссии по которому «классической» порождающей фонологии, допускавшей весьма абстрактные «промежуточные» репрезентации, противостояло, скажем, такое течение, как «естественная порождающая фонология» [Hooper 1976]. С другой стороны, развитие абстрактных репрезентационных теорий привело к такому важнейшему достижению, как появление автосегментной фонологии. В начале 1980-х годов с развитием лексической фонологии внимание исследователей обратилось к проблемам уровневых представлений и упорядочения правил, а также взаимодействия (мор)фонологии с морфологическими процессами, в первую очередь различными типами аффиксации. Хороший обзор этой проблематики дается в работе [Anderson 1985], ср. также [Кодзасов, Кривнова 1982].

Теория оптимальности также обязана своим появлением обсуждению такой «горячей темы», а именно дискуссии о так называемой просодической морфологии. К середине 1980-х годов важное место в фонологической литературе начали занимать формальные описания таких явлений, как редупликации, усечения и им подобные; с другой стороны, с развитием метрической теории слога деления и ударения привели к развитию теории «просодических»¹ единиц, которая оказалась весьма полезной для формулирования обобщений относительно упомянутых (мор)фонологических процессов. Весьма влиятельным оказалось понятие шаблона, то есть образца, к которому «должно было» прийти промежуточное представление. Так, для санскритской редупликации *du-druv-* в работе [Steriade 1988] предлагался следующий механизм порождения:

Исходные данные: *druv*, редупликация, шаблон CV

Правило 1 (полная редупликация): *druv* → *druv-druv*

Правило 2 (приспособление к шаблону):

Правило 2.1 (устранение кластера в приступе слога): *druv-druv* → *duv-druv*

Правило 2.2 (устранение коды): *duv-druv* → *du-druv*

Очевидно было, что объяснительная сила такой теории весьма невелика: в первую очередь неясно, почему шаблоны имеют ту форму, которая действительно встречается (а набор таких форм оказался совсем невелик). Однако ценность этого подхода заключалась в том, что в его рамках внимание исследователей было обращено не только (и, может быть, даже не столько) на абстрактные промежуточные стадии деривации, сколько на реальные ограничения, действие которых можно непосредственно наблюдать в засвидетельствованных формах.

Изначально теория оптимальности развивалась как раз для рассмотрения этих проблем (в чем можно убедиться, рассмотрев примеры в основополагающей работе [Prince,

¹ Точнее, вероятно, было бы называть их «супрасегментными»: речь идет о просодии в некотором широком смысле, как о комплексе явлений, затрагивающих единицы, отличные от сегмента (мора, слог, стопа и т.д.).

Smolensky 1993/2004]. Об этой «родовой травме» стоит помнить при чтении настоящего обзора. С ней связано не только то, что большая часть работ в рамках ТО, особенно ранних, посвящена таким, казалось бы, маргинальным вопросам, как форма редупликанта или механизм усечения в уменьшительных именах. Вероятно, именно своей истории ТО обязана тем, что в ее рамках только сейчас ставятся или не ставятся вовсе² проблемы, центральные для других теорий.

Итак, перейдем к изложению основных положений самой теории оптимальности. Цели, которые преследовали ее создатели, можно сформулировать следующим образом: создать такую теорию, которая при минимальном числе произвольных исходных допущений позволяла бы делать максимально точные предсказания относительно форм, возможных в человеческом языке; при этом к теории предъявляются дополнительные требования. Помимо стандартного набора (экономичность, формализованность, «простота» и др.) важную роль играла возможность моделировать усвоение языка. Речь здесь идет в первую очередь о том, чтобы уметь по набору исходных данных построить «правильную» (порождающую формы, в идеале полностью соответствующие наблюдаемым) грамматику, т.е. набор правил в соответствующем формализме. Напомним, согласно постулатам генеративного направления именно «умение» данной теории порождать все правильные формы, не порождая неправильных, является критерием так называемой эмпирической адекватности модели. При этом объяснительная адекватность, то есть соответствие модели реальным процессам в сознании говорящего, требует также моделирования других когнитивных процессов, к которым несомненно относится и усвоение языка.

Общий алгоритм, описывающий порождение форм в теории оптимальности, таков. На вход подается некая форма, которая берется из словаря. Эта форма является «глубинным» представлением (*underlying representation*), или просто «входом» (*input*). Для каждой такой формы порождается множество потенциальных кандидатов (т.е. тех форм, которые могут оказаться соответствующими данному входу на уровне поверхностных представлений). Это множество является потенциально бесконечным, то есть в него входят все мыслимые фонологические представления, причем это верно для всех языков. Иными словами, и для русского глубинного представления [knʲiga], и для нидерландского [buk] в качестве кандидата рассматривается, скажем, последовательность [inlwadi] – другое дело, что по целому ряду причин этот кандидат не является «победителем». Модуль порождения кандидатов называется Gen.

Следующий модуль после Gen – Eval, или модуль оценки. Его можно представить как функцию от двух аргументов, один из которых – пара <вход, кандидат_n>, а второй – кортеж из так называемых ограничений (*constraints*). На ограничениях определено транзитивное отношение доминирования (*domination*), задающее их порядок (в идеале – совершенный, то есть такой, в котором каждое ограничение доминирует над последующим, но подавлено предыдущим, а ситуация типа $W > X, Y > Z$ ³ запрещена).

Для каждой пары <вход, кандидат_n> и каждого ограничения можно сказать, сколько раз данная пара это ограничение нарушает. После оценки каждого кандидата проводится их сравнение. Чтобы данный кандидат оказался «победителем» (т.е. предсказываемой поверхностной формой) необходимо и достаточно, чтобы он наименьшее число раз нарушал наиболее высоко ранжированное ограничение (если два кандидата равное число нарушают некое ограничение, решение принимается на основании тех, что ранжиро-

² Можно было бы сказать: «замалчиваются». Но это, вероятно, будет не вполне справедливо: дело как раз в том, что фонология развивается вокруг своеобразных «точек роста», и, если какая-то проблема (быть может, весьма существенная) не попадает в зону внимания исследователей, вовсе не обязательно подозревать здесь умысел.

³ W доминирует над X и над Y, X и Y оба доминируют над Z (тогда в силу транзитивности W доминирует над Z), но неизвестно, $X > Y$ или $Y > X$ (легко убедиться, что обе гипотезы непротиворечивы).

ваны ниже). Здесь наиболее существенно, что общее количество нарушений не столь важно, как ранг самого высокого из тех ограничений, которые данная форма нарушает: даже если кандидат F1 нарушает низко ранжированное ограничение C1, скажем, семь раз, но не нарушает высоко ранжированное ограничение C2, он является оптимальным по сравнению с кандидатом F2, нарушившим C2 и вовсе не нарушившим C1. Эта логика называется с т р о г о й д о м и н а ц и е й (strict domination).

Обычно процедура выбора представляется в виде специальной таблицы (t a b l e a u), в наглядном виде представляющей этот процесс. В левом верхнем углу традиционно пишется вход. Каждая колонка соответствует одному ограничению: расположены они слева направо в порядке убывания ранга. Строки соответствуют кандидатам. Каждое нарушение одного ограничения кандидатом маркируется знаком * в соответствующей клетке: здесь существенно, что нарушений может быть и больше одного, к примеру, форма с двумя закрытыми слогами дважды нарушает ограничение NoCoda («коды запрещены» = «закрытые слоги запрещены»).

Согласно логике строгого доминирования, если некий кандидат нарушает какое-либо высоко ранжированное ограничение (и есть кандидаты, это ограничение не нарушающие), он выбывает из рассмотрения, так как количество нарушений не столь важных ограничений несущественно. В таблице это обозначается затенением всех клеток «низких» ограничений.

Приведем пример такой таблицы (из работы [Akinlabi 1996]), иллюстрирующей «мобильную» лабиализацию в одном из диалектов эфиосемитского языка гураге. Согласно А. Акинлаби, форма типа *näk^wäs* ‘укусил его’ порождается из соположения глубинных представлений *näk^wäs* ‘укусил’ и [round] (фонологический признак) ‘3sg объекта’. Также в языке гураге действуют следующие ограничения:

- *COR / LAB «если сегмент имеет спецификацию [coronal], он не может иметь спецификацию [labial]», что в терминах сегментного набора гураге означает запрет на лабиализованные переднеязычные;
- PARSEMORPH «все морфемы, присутствующие в глубинном представлении, должны иметь ненулевое соответствие в поверхностном представлении»;
- ALIGN-R[ROUND] «признак [round] должен реализоваться на правом краю морфологической основы».

Таблица 1

<i>näkäs</i> , [round]	*COR/LAB	PARSEMORPH	ALIGN-R[ROUND]
<i>näkäs</i>		!*	
<i>ɛä</i> <i>näk^wäs</i>			*
<i>näkäs^w</i>	!*		

Ранжирование ограничений в этом примере таково: *COR/LAB ≧ PARSEMORPH ≧ ALIGN-R[ROUND]. Как можно видеть, хотя оптимальный кандидат *näk^wäs* нарушает ограничение Align-R (так как [round] не реализуется на правом краю основы), этот факт уже не является существенным для процедуры выбора: два других кандидата уже не могут быть рассмотрены в силу того, что нарушают более высоко ранжированные ограничения, – один не имеет никакого лабиализованного сегмента, то есть в нем нет соответствия морфеме, состоящей из признака [round], а другой содержит запрещенный ограничением *COR/LAB лабиализованный переднеязычный сегмент.

Здесь возникает законный вопрос: если потенциальное множество кандидатов бесконечно, как исследователь выбирает те, которые он рассматривает? Очевидно, что здесь открывается обширное пространство для произвола, на что неоднократно указывалось в литературе: ведь в конечном счете множество реально рассматриваемых кандидатов в

таком случае определяется лишь доброй волей лингвиста. Как показывает опыт, зачастую в множество рассматриваемых кандидатов действительно могут не попадать те, которые могут опровергнуть предложенную теорию: разбору одного такого случая посвящена работа [Karttunen 2006], где, впрочем, предлагается основанный на теории конечных автоматов механизм порождения множества кандидатов.

Нулевая гипотеза теории оптимальности звучит так: различия между языками сводятся к разнице в ранжировании ограничений, набор которых универсален. Одно из важных дополнительных предположений заключается в том, что эти различия не могут объясняться путем введения дополнительных произвольных ограничений на структуру входов (глубинных форм). Этот постулат, имеющий в литературе название «богатство базы» (Richness of the Base), весьма важен для понимания разницы между ТО и многими деривационными подходами, в первую очередь классической порождающей фонологии. В статье [McCarthy 2004a], где подробно разбирается важность «богатства базы», эта разница формулируется примерно так: если в порождающей фонологии бремя доказательства несла не только деривация (то есть собственно правила и порядок их применения), но и исходная репрезентация (формализуемая, например, в виде ограничений на морфемную структуру (Morpheme Structure Constraints – MFCs)), то в ТО – по предположению – для объяснения различий между языками достаточно указать на разницу в деривации (порядке ограничений). Предполагается, что *ceteris paribus* (т.е. при равной объяснительной адекватности и т.д.) ТО лучше удовлетворяет условию бритвы Оккама: она не требует накладывать специальные условия на глубинные представления. Это значит, что вопрос о природе глубинных представлений вовсе не относится к компетенции лингвистической теории, а тогда в связи с ним нельзя сделать «лишних» допущений.

Практическое следствие «богатства базы», согласно цитированной статье Дж. Маккарти, должно заключаться в следующем. Коль скоро мы не имеем права *a priori* предполагать, что поверхностная форма выглядит именно так, как нужно, чтобы наша гипотеза оказалась верной, необходимо, проверяя свое предположение, провести дополнительную работу. Работа эта заключается в том, чтобы убедиться: если предположить другую глубинную форму, предложенное ранжирование ограничений не предскажет неверную форму. Если возвратиться к примеру из гугаге выше, это значит, что предлагаемое ранжирование не должно, исходя из глубинной формы, скажем, **nākāsw* (а не *nākās*, [ground]), предсказать существование невозможной формы **nākās*^{*}. Стоит, однако, отметить, что, к сожалению, такая проверка проводится далеко не всегда (но эту претензию стоит предъявлять, конечно, не самой теории, а автору исследования). С другой стороны, сам постулат «богатства базы» в последнее время подвергается обоснованной критике: в первую очередь потому, что, по всей видимости, достаточно ограничить мощность множества предсказываемых языков путем одного переранжирования ограничений невозможно.

Здесь нужно подчеркнуть одно свойство ТО, которое ее сторонники считают весьма существенным: по своей природе теория оптимальности – теория типологическая. Во-первых, это следует понимать в достаточно прямолинейном смысле: ТО может (в идеале) предсказать все разнообразие возможных языков (если предположить, что число ограничений конечно, очевидно, что и число их перестановок конечно, а каждая перестановка задаст возможную грамматику). Мощность множества этих грамматик, конечно, очень велика (на что упирают критики), но она все же меньше мощности множества тех грамматик, что теоретически способна предсказать деривационная теория, вооруженная такими мощными инструментами, как сам механизм правила⁴, явное задание их порядка (extrinsic ordering) и потенциально неограниченное число уровней деривации (и это, по заверениям сторонников ТО, уже достижение).

⁴ Ничто не запрещает предположить такое правило, как, скажем, $b \rightarrow x/V_t'$, хотя типологически оно выглядит в высшей степени неестественно.

Прежде чем обратиться ко второму аспекту «типологичности» ТО, необходимо рассмотреть виды ограничений и способы их обнаружения: очевидно, что именно число ограничений определяет мощность множества предсказываемых языков, и в то же время именно здесь открыт наибольший простор для произвола исследователя. Действительно, если гипотеза хороша всем, кроме того, что реально наблюдаемая форма в ее рамках является не оптимальной, но «второй по оптимальности», велик соблазн *ad hoc* ввести некое ограничение, единственный смысл которого: запретить неудобный кандидат. При этом набор ограничений, напомним, универсален, то есть ограничение, постулируемое для английского языка, будет присутствовать (может быть, с очень низким рангом) и в варлпири, и в арабском, и в русском.

В «классической» ТО все ограничения делятся на две большие группы: ограничения на соотношение между глубинной и поверхностной формой (*faithfulness constraints*, ограничения «преданности»⁵) и ограничения на структуру поверхностной формы (*markedness constraints*, ограничения маркированности).

Ограничения преданности в первую очередь отвечают за то, чтобы глубинная и поверхностная форма были «достаточно похожи» (необходимость такой похожести в рамках ТО является аксиомой, хотя, строго говоря, этот вопрос может требовать дополнительного обсуждения). Сюда входит как «похожесть» (а лучше – идентичность) сегментов и их признаковой структуры, так и, к примеру, постулат «постоянного выражения [морфологических категорий]» (*consistency of exponence*), который гласит, что сегменты⁶ не могут при переходе от глубинной к поверхностной форме менять свою морфологическую принадлежность [*van Oostendorp 2007*].

Что касается формализации ограничений преданности, то в разных версиях ТО существуют разные мнения на этот счет. Изначально А. Принс и Дж. Маккарти предполагали, что глубинная и поверхностная форма связаны отношением вхождения (глубинная форма «содержалась»⁷ в поверхностной). Позже в работе [*McCarthy, Prince 1995*] была предложена так называемая теория соответствия (*correspondence theory*), предполагавшая между элементами входа и выхода отношение соответствия (подробнее об этом см. ниже). В последнее время в некоторых работах (например, у М. ван Остендорпа) можно встретить предложения вернуться к теории вхождения, в частности для описания нетривиальных чередований, ограниченных теми или иными морфологическими соображениями (которые никогда не были для ТО сильным местом).

Однако гораздо более яркую особенность ТО представляет предлагаемое в ее рамках понимание маркированности. Мы не будем здесь обсуждать развитие понятия маркированности в разных теориях, отметим лишь, что в рамках ТО она понимается не только чисто формально (как у Трубецкого – наличие положительного значения признака в рамках привативной оппозиции) – в понятие «маркированности» вкладывается достаточно серьезное содержание. В частности, ТО предсказывает, что «маркированные» феномены будут встречаться в языках мира реже немаркированных. Однако «диагностический признак» маркированности – отнюдь не в этом.

Маркированность играет в ТО роль «противовеса» преданности. В частности, в статье [*Moroton 1994*] показано, что практически все фонологические процессы (за исключением сравнительно редких вроде метатез) связаны исключительно либо с выигрышем в преданности (то есть созданию «нехороших» с точки зрения условий на поверхностное представление структур, лучше удовлетворяющих требованиям близости глубинной и поверхностной форм), либо с созданием менее маркированных структур, отклоняющихся от глубинных.

Таким образом, ТО является инструментом непосредственной формализации давно известного в лингвистике конфликта необходимости различать формы (потребность

⁵ Этот русский вариант термина, кажущийся нам удачным, принадлежит С.Г. Татевосову.

⁶ И шире – автосегменты, неассоциированные признаки и пр.

⁷ Contained. Отсюда название этой теории – Containment theory.

слушающего) и создавать «несложные» структуры (потребность говорящего). Как справедливо замечает М. Хаспельмат [Haspelmath (в печати)], в этом смысле ТО занимает промежуточную позицию между формальным («генеративистским») и функционалистским подходом.

Тот факт, что маркированность является противовесом преданности, подводит нас к способу формулирования связанных с ней ограничений. Напомним, что грамматика каждого языка соответствует некоторому ранжированию универсальных ограничений, а разные грамматики получаются путем смены рангов. Тогда очевидно следующее: если мы предположим, что в универсальном наборе существует некое ограничение C1, дающее в языке L (при ранжировании R) определенный эффект, то, изменив R, мы получим новый язык L', предсказуемо отличный от L. Если окажется, что таких языков L' не существует, вероятнее всего, ограничение C1 не должно использоваться при анализе, и vice versa.

Так, например, в работе [Beckman 1998] обосновывается необходимость специального ограничения MAX(V̄) «Каждый долгий гласный в глубинной форме имеет соответствие в поверхностной», отличного от более общих Max(V) («сохраняются все гласные независимо от долготы») или MaxSeg («сохраняются все сегменты»). Обосновывается это следующим образом.

Предположим, что ограничение MAX(V̄) существует. Оно может быть активно в фонологических процессах, управляющих удалением гласных (синкопа/апокопа). Сами эти процессы связаны с соображениями маркированности, в первую очередь маркированности метрической структуры: синкопа и/или апокопа создают «лучшие» (более короткие) стопы и/или слова. Ограничения, отвечающие за эти процессы, кратко обозначим FtStruc (здесь неважно, какой в точности механизм действует, если он действительно приводит к ожидаемым эффектам). Очевидно, в языках без апокопы FtStruc ранжировано достаточно низко – по крайней мере ниже Max(V), что уже достаточно для отсутствия соответствующих процессов в фонологии данного языка.

Теперь проведем мысленный эксперимент и посмотрим, какие языки предсказываются различными переранжированиями этих трех ограничений. Пусть FtStruc требует синкопы второго гласного в последовательности трех открытых слогов. Сравним поведение гипотетических слов с долгим и кратким гласным во втором слоге при разных ранжированиях (табл. 2):

Таблица 2

$$FtStruc \geq Max(V) \geq Max(\bar{V})$$

tatāta	*FtSTRUC	Max(V)	MAX(V̄)
tatta		*	*
tatata	!*		*
tatāta	!*		
tatata	*FtSTRUC	Max(V)	MAX(V̄)
tatta		*	*
tatata	!*		
tatāta	!*		

Как легко видеть, если маркированность ранжирована выше преданности для всех типов гласных, процесс, за который «отвечает» высоко ранжированное ограничение, проходит во всех формах – и с кратким, и с долгим гласным. Такой естественный язык вполне представим. Легко убедиться, что перестановка низко ранжированных ограничений ($FTSTRUC \gg MAX(\bar{V}) \gg MAX(V)$) в этом случае никакого видимого эффекта иметь не будет (однако см. ниже о «возникновении немаркированности»).

Теперь рассмотрим тот случай, когда ограничение преданности $MAX(V)$ ранжировано выше маркированности $FTSTRUC$, а та, в свою очередь – выше $MAX(\bar{V})$ – см. табл. 3.

Таблица 3

MAX(V) ≫ FTSTRUC ≫ MAX($\bar{V}$)			
tatāta	MAX(V)	*FTSTRUC	MAX($\bar{V}$)
tatta	!*		*
✎ tatata		*	
✎ tatāta		*	

tatata	MAX(V)	*FTSTRUC	MAX($\bar{V}$)
tatta	!*		
✎ tatata		*	
tatāta		*	

Здесь сразу выпадает из рассмотрения кандидат, удовлетворяющий требованиям маркированности и потому отклоняющийся от глубинной формы; что касается долготы гласного, то ограничение $MAX(\bar{V})$ попросту не может играть никакой роли, так как оно не говорит ничего о кратких гласных и не сообщает, каким именно должно быть соответствие долгову гласному (напомним, это ограничение требует лишь, чтобы такое соответствие существовало). Таким образом, этих трех ограничений недостаточно, чтобы решить, какой из кандидатов окажется оптимальным. Здесь нужно привлечь данные о ранжировании других ограничений (в данном случае – DER-μ, «каждая мора в поверхностном представлении имеет соответствие в глубинном», проще говоря – «вставка мор запрещена») в интересующем нас языке⁸. Таким образом, приведенное ранжирование задает язык, в котором синкопа вовсе не происходит (что, конечно, типологически вполне обычно). Опять-таки легко убедиться, что ранжирования $MAX(V) \gg MAX(\bar{V}) \gg FTSTRUC$ и $MAX(\bar{V}) \gg MAX(V) \gg FTSTRUC$ (иначе говоря, все, где маркированность ранжирована ниже «общей» преданности) приводят к тому же результату.

Наконец, самый интересный случай: такое ранжирование, при котором «позиционная преданность» (т.е. ограничение преданности, касающееся ограниченного класса объектов, в нашем случае долгих гласных) ранжирована выше маркированности (см. табл. 4).

⁸ Логически представимо, чтобы ограничение DER-μ по той или иной причине было неактивно: в таком случае принято считать, что две формы находятся в отношении свободного варьирования.

$$\text{MAX}(\bar{V}) \gg \text{FTSTRUC} \gg \text{MAX}(V)$$

<i>tatāta</i>	$\text{MAX}(\bar{V})$	*FTSTRUC	$\text{MAX}(V)$
<i>tatta</i>	!* ⁹		*
<i>tatata</i>		*	
<i>tatāta</i>		*	

<i>tatata</i>	$\text{MAX}(\bar{V})$	*FTSTRUC	$\text{MAX}(V)$
<i>tatta</i>			*
<i>tatata</i>		!* ⁹	
<i>tatāta</i>		!* ⁹	

Это ранжирование приводит к ожидаемому результату: «позиционная преданность» «спасает» долгие гласные от синкопы, притом что «в общем случае» (при кратких гласных) синкопа действительно происходит, и в таких языках, как показала Дж. Бекман, недостатка также нет.

Таким образом, переранжирование небольшого числа ограничений предсказывает как раз то типологическое разнообразие, которое наблюдается в языках мира, и это серьезный аргумент в пользу того, что именно эти ограничения являются частью универсального набора.

Как можно убедиться, рассмотренный нами механизм имеет весьма широкую применимость. В рамках ТО возможно несколько вариантов взаимодействия маркированности и преданности⁹.

1. Маркированность ранжирована выше преданности: эффект, требуемый данным ограничением маркированности, наблюдается во всех поверхностных формах данного языка (в литературе такие явления иногда называются *surface-true generalizations*); см. табл. 2.
2. Преданность ранжирована выше маркированности: эффект, требуемый ограничением маркированности, при отсутствии сторонних факторов не наблюдается; см. табл. 3.

Кроме того, существует два смешанных случая, один из которых мы проиллюстрировали в табл. 4. Пусть F – ограничение преданности, M – ограничение маркированности, и эти ограничения каким-то образом взаимодействуют.

3. $F_1 \gg M \gg F_2$, причем F_1 действует в подмножестве контекстов¹⁰ действия F_2 (см. табл. 4). В этом случае эффект, требуемый ограничением M , будет наблюдаться только в тех контекстах, где действует F_2 , но не F_1 – это и называется «позицион-

⁹ Напомним, речь не идет о «глобальном ранжировании» типа «все ограничения маркированности ранжированы выше всех ограничений преданности»: каждый раз рассматривается несколько ограничений, каждое из которых относится к одному из двух классов.

¹⁰ В противном случае мы попросту имеем дело с частным случаем того, что описано в пункте 2.

ной преданностью», то есть сохранением глубинной формы в определенных позициях.

4. $F_1 \gg M \gg F_2$, причем есть контексты, где применимо F_2 , но не F_1 . Этот случай имеет специальное название – «возникновение немаркированности» (the emergence of the unmarked, TETU; см. [McCarthy, Prince 1994]). В этом случае эффект, требуемый ограничением M , наблюдаться почти никогда не будет, так как ограничение F_1 запрещает отклоняться от глубинной формы. Если же мы предположим, что в каком-то контексте ограничение F_1 не действует, то нам останется рассмотреть ранжирование более низких ограничений, а там мы наблюдаем ситуацию, описанную в пункте 1: требования маркированности доминируют над преданностью. Таким образом, низко ранжированные ограничения вовсе не обязательно являются полностью неактивными: напротив, они могут оказывать существенное влияние на поверхностные формы.

Наиболее яркие примеры «возникновения немаркированного» известны из различных типов редупликаций, в которых достаточно «неудобные» структуры допустимы в базе, но не в редупликанте. Предполагается, что в базе действуют ограничения преданности (так как основа имеет лексическое наполнение), а в редупликанте – нет (так как в нем нет лексически заданных сегментов, все они являются копией соответствующего материала из базы). Известен пример из языка нутка, где редупликант имеет форму CV. В нутка действует следующее ранжирование: MAX-IO $\gg$ NOCODA $\gg$ MAX-BR, где MAX-IO требует, чтобы все сегменты входа имели соответствие в выходе, NOCODA запрещает закрытые слоги, а MAX-BR требует, чтобы все элементы базы копировались в редупликант (RED – пустая с точки зрения сегментного состава редупликативная морфема), см. табл. 5.

Таблица 5

RED-čims' i:ħ	MAX-IO	NOCODA	MAX-BR
či-čims' i:ħ		**	****
čim-čims' i:ħ		**!*	***
čims' i:-čims' i:ħ		**!*	*
čims' i:ħ-čims' i:ħ		**!*	

Легко убедиться, что в нередуплицированной форме ограничение NOCODA неактивно, так как любое усечение нарушало бы ранжированное выше MAX-IO¹¹.

Вернемся к вопросу об оправданности введения тех или иных ограничений. Поскольку известно, к каким эффектам приводят те или иные виды переранжирования, мы можем, выдвигая гипотезу о существовании в универсальном наборе того или иного ограничения, проверить ее типологически. В самом деле, если бы, скажем, вместо MAX($\bar{V}$) мы попытались ввести ограничение Max($\bar{V}$)¹², «защищающее» краткие гласные, мы ожидали бы существования таких языков, где процессы вроде синкопы затрагивали бы долгие, но не краткие гласные, а такие языки не засвидетельствованы.

Итак, в ТО существует два типа ограничений: ограничения преданности, обеспечивающие близость глубинных и поверхностных форм, и ограничения маркированности, формализующие требования к структуре форм в фонологическом представлении.

¹¹ Очевидно, DEP-IO (запрет на вставку сегментов) также ранжировано достаточно высоко, иначе имелись бы формы без закрытых слогов типа *čimis' i:ħi.

¹² Это рассуждение приводится в [McCarthy 1998].

Ограничения маркированности обыкновенно имеют вид запретов на «неудобные» структуры. Строго говоря, вопрос о том, почему та или иная структура является «маркированной» (в смысле ТО), чисто формален – структура является таковой, если ее запрещает некоторое универсальное ограничение, а набор ограничений выводится, как мы видели, исключительно на основе эмпирических данных. Тем не менее ясно, что на деле ограничения маркированности позволяют формализовать как раз те свойства, которые в других фонологических теориях часто приходится постулировать как присущие маркированным (в традиционном смысле) единицам. Иначе говоря, эти свойства вытекают из более общих свойств всей грамматики (системы ранжированных ограничений), а не из частных аксиом.

Известный пример такого рода явления – факториальная типология слога. Еще Якобсон показал, что типов позволенных слогов в языках мира всего четыре:

1. (C)V(C) (английский, русский);
2. CV(C) (йокуц);
3. CV (сенуфо);
4. (C)V (гавайский).

При этом логически таких типов возможно девять, то есть пять типов разрешенной слоговой структуры не засвидетельствовано ни в одном языке (например, не существует языков, в которых слоги обязательно закрыты, но необязательно прикрыты ((C)VC).

В [Prince, Smolensky 1993/2002] типология слога разобрана весьма подробно. Там показано, что если принять три основных ограничения: преданность (F)¹³, NoCoda (закрытые слоги запрещены) и Onset (неприкрытые слоги запрещены), то их факториальная типология предсказывает в точности засвидетельствованное варьирование (см. табл. 6).

Таблица 6

Ранжирование	Тип языка
F ≧ Onset, NoCoda	(C)V(C)
Onset ≧ F ≧ NoCoda	CV(C)
Onset, NoCoda ≧ F	CV
NoCoda ≧ F ≧ Onset	(C)V

При этом, как отмечают Принс и Смоленский, есть существенная разница между традиционным анализом (согласно которому принципы наподобие NoCoda и Onset в соответствующих языках не подавлены, а попросту отсутствуют). Так, для языков, в которых ограничение Onset не является недоминированным (то есть существуют и слоги без приступов), его действие все же можно наблюдать: например, именно оно отвечает за тот факт, что последовательность CVCV делится на слоги как .CV.CV. (а не, скажем, логически вполне возможное .CVC.V.).

Таким образом, теперь типологию слогов в языках мира можно задавать не списком, а как следствие определенного ранжирования универсальных элементов. Остается, однако, вопрос: почему нельзя сделать того же с помощью деривационных теорий, основанных на правилах?

В работе [Hammond 1997] на этом примере указано на очень важное свойство ТО, которое является едва ли не главной отличительной особенностью этой теории: деривационные теории ориентированы на структуру *глубинного* представления, а для ТО гораздо важнее *поверхностная* форма. Вообще говоря, структура *глубинного*

¹³ У Принса и Смоленского типология более дробна, но для наших целей достаточно и этого.

представления в ТО вовсе не важна, при условии, что грамматика адекватна данным. Разобранный выше постулат «богатства базы» можно сформулировать так: «О глубинных формах нельзя сказать ничего существенного». Для деривационных теорий, напротив, (не)применимость правила и его действие определяются свойствами входа («более глубинного» представления) и контекста (существующего на том же уровне деривации, что и вход).

При этом очень многие феномены поддаются гораздо более простому и экономному описанию при ориентации именно на поверхностную форму. В деривационной фонологии эта проблема получила название проблемы заговоров (conspiracies): деривационный анализ требует введения нескольких разнотипных правил, приводящих к появлению некоторой весьма просто описываемой структуры. Проблема здесь состоит в том, что «простое», интуитивное объяснение («существует определенное ограничение на поверхностную структуру») в рамках деривационного подхода оказывается несущественным следствием применения системы никак не связанных правил, зачастую весьма сложной. В [Kager 1999] это формулируется так: «...ТО предсказывает, что некое ограничение маркированности может приводить к различным типам структурных изменений, в зависимости от взаимодействия с ограничениями преданности. Таким образом, в разных языках могут применяться разные “стратегии исправления” для достижения одной и той же цели в поверхностной форме... Напротив, деривационная теория не может предсказать подобного функционального единства процессов, поскольку в ней не существует формального способа представить понятие “цели” фонологического правила»¹⁴. Так, например, в языке кваньяма (группа банту, R20 по Гасри) запрещены сочетания «носовой + глухой смычный» и существуют два способа их избежать: озвончение глухого (как в *sitamba* «печать», из англ. *stamp*) и прогрессивная ассимиляция (как в *onana* «теленки», морфологически *on-tana*). Очевидно, что в рамках деривационной теории потребовалось бы два никак не связанных правила, что только затемнило бы достаточно ясную картину (подробнее об этом случае см. [Kager 1999: 83–86]). Еще один пример подобного «заговора» приведен в [Hammond 1997: 46–50] и связан с синкопой в разговорном английском: согласно Хэммонду, деривационный анализ требует постулирования трех правил, в то время как простейшее описание (достаточно легко формализуемое в ТО) связано с тем, какая структура представлена в поверхностной форме (проще говоря, важнее не то, что было до синкопы, а то, к чему она привела).

Еще один аспект, связанный с ориентацией на поверхностную форму, – решение так называемой «проблемы дублирования» (duplication problem), связанной с существованием в порождающей фонологии не только собственно деривационных правил, но и «ограничений на структуру морфемы» (morpheme structure constraints). Эти ограничения были введены как способ формализации статических обобщений, верных для лексических представлений слов¹⁵. Например, в русском языке все гласные фонемы заднего ряда являются огубленными, а те, что не относятся к заднему ряду – неогубленными; это формализуется с помощью неполной спецификации признака огубленности в словаре и применением соответствующего «ограничения на структуру морфемы», т.е. правила, применяемого до начала собственно деривации: [-back 0 round] → [-back -round] и [+back 0 round] → [+back +round]. Существенно, что в рамках самой деривации (то есть применения тех правил, существование которых в порождающей фонологии устанавливается на основе чередований) такие правила не участвуют. Похожее ограничение существует и в австронезийском языке чаморро: в нем имеется шесть гласных (*i, u, e, o, ä, a*), причем

¹⁴ «...OT predicts that a markedness constraint may trigger various types of structural changes, depending on its interaction with faithfulness constraints. Different languages should therefore pursue different “repair strategies” in attaining identical output goals... In contrast, a rule-based theory fails to make this prediction of the functional unity of processes because it has no formal means of expressing the notion of “output goal” of a phonological rule» [Kager 1999: 55–56].

¹⁵ Изложение этой проблемы основано на [Kenstowicz, Kisseberth 1979]; подробнее см. также [Kenstowicz, Kisseberth 1977].

в подсистеме низких гласных признак огубленности также не является дистинктивным; следовательно, для чаморро имеет смысл постулировать похожие ограничения на морфемную структуру. Однако в чаморро существует и собственно фонологический процесс перегласовки (в первом слоге корня непредние гласные упередняются после переднего гласного): *u, o, a* → *i, e, ä*. Простейшая формулировка этого процесса в терминах правил ([+syll] → [-back] / [+syll -back]C_) приводит к неверным результатам, так как предсказывает существование в чаморро гласных *ö* и *ÿ*; следовательно, необходимо дополнительное правило (в деривации!), переводящее низкие передние огубленные гласные в неогубленные. Нетрудно заметить, что это правило будет в точности повторять уже примененное ограничение на структуру морфемы. Нетрудно заметить и функциональное единство этих двух отдельных (хотя и идентичных процессов). Таким образом, одна и та же информация о поверхностной форме кодируется в порождающей фонологии двумя отдельными правилами. Очевидно, что в ТО эта проблема не встает, так как (в идеале) нужные формы обеспечиваются соответствующим ранжированием ограничений маркированности (по сути являющихся формализацией эксплицитно заявленного принципа: в нашем случае «в чаморро нет передних огубленных гласных») и преданности.

В заключение следует заметить, что сам аппарат ТО приложим не только к поверхностным представлениям: с помощью подобных алгоритмов можно оценивать не только пары <вход, кандидат>, но и другие объекты: ср., например [Blumenfeld 2006], где алгоритм ТО применяется к оценке не самих форм, но тех процессов, которые необходимы, чтобы перейти от данного входа к данному выходу.

Наконец, еще один важный аспект классической теории оптимальности: ее принципиальная двухуровневость: в отличие от многоступенчатой деривации в порождающей или лексической фонологии в ТО предлагается считать, что для перехода от глубинной формы к поверхностной достаточно одного шага. Эта гипотеза является одной из наиболее обсуждаемых в литературе по ТО, и дать о ней сколько-нибудь полное представление в кратком обзоре невозможно. Тем не менее одну проблему в связи с этим обсудить все же необходимо: это проблема так называемой «непрозрачности» (opacity).

В деривационных теориях наподобие лексической фонологии большую роль играли вопросы, связанные с взаимодействием правил и контекстов их применения. Различалось несколько типов такого взаимодействия: *недоприменение* (underapplication) – некоторое правило не применяется, хотя в поверхностной форме существует контекст для его применения – и *переприменение* (overapplication) – некоторое правило применяется, хотя контекста для его применения все же нет¹⁶. Известный пример непрозрачности с переприменением – взаимодействие эпентезы и удаления велярного в турецком языке. В турецком сочетания согласных в конце слова разбиваются гласным, участвующим в гармонии: [baʃ-m] ‘голова-его’ → [baʃɪm], [jel-m] ‘ветер-мой’ → [jelim]. Сегмент [k] в интервокальной позиции недопустим и потому чередуется с нулем: [ajak-I] ‘нога-моя’ → [ajaɪ]. При взаимодействии этих правил вначале применяется эпентеза: [ajak-m] → [ajaɪm]. Деривационный анализ сравнительно прост (хотя и повторяет диахронические переходы – впрочем, это давно известная проблема порождающей фонологии), но при этом для него ключевое значение имеет существование промежуточных стадий, ведь только на одной из них и возможно применение правила удаления [k].

Пример непрозрачности с недоприменением можно найти в одном из диалектов языка науатль, где взаимодействуют правила апокопы (необязательной) и оглушения аппроксимантов в конце слова. Оглушение применяется до апокопы, и, следовательно, если апокопа приводит к тому, что некий аппроксимант оказывается в конце слова, оглушения уже не происходит: ср. [ˈtajoː!] ‘очищенная кукуруза’ (лексически *tajoːl*) и [ʃikaˈkili] ~ [ʃikaˈkɪl] ‘положи это туда’ (лексически *ʃikaˈkili*). Оказывается, что классиче-

¹⁶ Приводимые здесь примеры взяты из [Kager 1999: 372–400]; там же см. подробное обсуждение и библиографию.

ская ТО с ее ориентацией исключительно на поверхностную форму никак не может решить этих проблем: так, в случае непрозрачности с переприменением множество нарушений, связанных с неверно предсказываемым кандидатом, будет собственным подмножеством множества тех, что связаны с реально наблюдаемым: таким образом, при любом ранжировании ограничений реально наблюдаемый кандидат не сможет выиграть (так как по всем релевантным ограничениям будет «ничья», но при этом у него будут еще и «лишние» нарушения, которых не будет у неверно предсказываемой формы).

Предложено несколько способов решения этой проблемы:

- «Двухуровневая маркированность» (например [Orgun 1996] – *correspondence and identity constraints*): ограничения маркированности могут не только запрещать определенные поверхностные формы, но и иметь доступ к информации о глубинной форме.
- Многоуровневая деривация, оптимальная лексическая фонология [Bermúdez-Otero (в печати)]. Здесь алгоритмы ТО применяются на нескольких уровнях, совмещая преимущества двухуровневого подхода с объяснительной силой деривационных теорий.
- «Теория симпатии» [McCarthy 1998]. В этой модели оценивается соответствие поверхностной формы не только глубинной, но и одному из (невыигрывающих) кандидатов, выбираемому специальным образом. Хотя здесь сохраняются основные принципы ТО, резко увеличено число оцениваемых соответствий; к тому же не вполне понятно, чем содержательно эта модель отличается от сериалистских концепций.
- Локальная конъюнкция (например [Kirchner 1996]): существуют ограничения вида $C1 \ \& \ C2$ (причем всегда верно, что $C1 \ \& \ C2 \geq C1, C2$), нарушаемые тогда и только тогда, когда один и тот же элемент (скажем, сегмент), нарушает и $C1$, и $C2$. Несмотря на определенные успехи в сфере эмпирической адекватности, локальная конъюнкция нарушает логику строгой доминации и резко увеличивает множество теоретически мыслимых ограничений (что снижает объяснительную силу теории и к тому же имеет серьезные последствия для вопроса о том, возможно ли усвоить такую грамматику по поверхностным данным).
- «Сравнительная маркированность» [McCarthy 2003] – ограничения маркированности делятся на два класса: те, которые запрещают структуры, встречающиеся в глубинной форме, и те, которые «вводятся» данным кандидатом.
- «ТО с цепочками кандидатов» [McCarthy 2007] – еще одна псевдодеривационная теория, где кандидаты представляют собой цепочки, в которых каждое следующее звено более гармонично, чем предыдущее.

В целом, однако, следует признать, что общепринятого решения этих проблем в ТО пока не предложено.

Открытым остается и вопрос о природе набора ограничений. Помимо классических маркированности и преданности в литературе можно встретить предложения ввести такие типы, как ограничения на соответствие нескольких поверхностных форм (*output-output correspondence*), которые также применяются для решения проблем, связанных с непрозрачностью, ограничения «анти-преданности» (логическое отрицание преданности) [Alderete 1998] и др. Вопрос об оправданности введения таких ограничений до сих пор является дискуссионным.

Как мы уже упоминали выше, применимость алгоритма ТО не ограничена одной лишь фонологией: существуют и ее приложения к другим уровням языка, в частности синтаксису и морфологии (см., к примеру, соответствующие главы в [Archangeli, Langendoen 1997]). Круг проблем, рассматриваемых в рамках ТО в этих дисциплинах, весьма широк: от описания некоторых типов морфологических процессов (те же редупликации и усечения, «неестественные» правила), порядка морфем в словоформе или сочетаемости в словообразовании до сложных вопросов синтаксиса (порядок слов, связывание и

пр.). Неудивительно, что особенно тесно ТО интегрирована с репрезентационными теориями (не подразумевающими многоуровневой деривации) – в особенности это касается лексико-функциональной грамматики (LFG, ср. [Sells 2001]). В рамках OT-LFG ведется достаточно активная работа, однако нужно помнить, что в принципе алгоритмы ТО не требуют жесткой привязки к конкретным типам синтаксических репрезентаций: ср. анализ контроля PRO в английском языке в [Speas 1997], фактически не прибегающий к каким-либо репрезентациям, специфичным для конкретных теорий.

Тем не менее следует отметить, что поскольку объем работы в рамках ТО в области синтаксиса и морфологии все же меньше, чем в фонологии, в этих дисциплинах пока не разработана столь же подробная теория типов ограничений (и менее устоялся их набор), что несколько снижает привлекательность ТО в нынешнем виде как инструмента для предсказаний типологического варьирования. Более того, как отмечено в [Sells 2003], в некоторых версиях ТО архитектура синтаксиса резко отличается от архитектуры фонологии. Это ставит под серьезное сомнение применимость ТО как обобщенной теории грамматики: впрочем, сам Селлз критикует упомянутые подходы. Следует отметить так же и то, что, по замечанию М. Хаспельмата [Haspelmath (в печати)], многие исследователи, работающие в области морфологической и синтаксической типологии, используют ТО как технический прием, удобный инструмент представления своих теорий, вовсе не обязательно принимая все ее теоретические постулаты касательно универсальной грамматики.

Объем настоящего обзора не позволяет нам подробно обсуждать другие вопросы, связанные с ТО, – ниже мы лишь укажем на некоторые из них, приведя отсылки к некоторым важным работам, с которых можно начинать знакомство с соответствующими темами.

Итак, попробуем перечислить некоторые из основных вопросов, которые удалось поставить (необязательно решить) в рамках ТО.

- В какой мере типологическое варьирование выводимо из небольшого числа независимых исходных допущений? Несомненно, такая постановка вопроса отнюдь не является новой со времен теории принципов и параметров, но именно ТО – по крайней мере в фонологии – позволила совершенно по-новому подойти к этой проблеме. Как мы видели выше, теория оптимальности позволяет «списать» на независимые принципы очень многое из того, что ранее приходилось постулировать как данность: ср. обсуждение структуры слога, где факториальная типология вытекает из ранжирования ограничений, наблюдаемых и в других фрагментах грамматики. В рамках ТО, в частности, удалось решить проблему постулирования шаблонов в просодической морфологии. В целом можно сказать, что одно из главных достижений теории оптимальности – отказ от аксиоматического постулирования различных феноменов, которые ранее могли быть заданы только таким образом.
- Как можно усвоить грамматическую (фонологическую) структуру языка? Можно ли это сделать, если она устроена в соответствии с принципами ТО? В рамках теории оптимальности получены интересные результаты в области математической лингвистики, в частности, разработаны алгоритмы, позволяющие по набору форм строить порождающую их грамматику (см. [Tesar, Smolensky 2000; Boersma, Hayes 2001]; ср. также [Keller, Asudeh 2002], соответствующий раздел в [Kager 1999]), доказано, что эти алгоритмы всегда приводят к построению грамматики. В рамках ТО вообще удалось поставить целый ряд таких вопросов о математических свойствах лингвистических систем, которые в других теориях даже не задавались – и в то же время имеют непосредственное отношение к насущным проблемам, в первую очередь к усвоению языка и к разрешимости проблем анализа и порождения языка (что имеет уже прямое практическое применение).
- Сколько уровней деривации необходимо в лингвистической теории? Этот вопрос имеет как теоретические (ср. выше о проблеме «непрозрачности»), так и практические следствия (в частности, признание многоуровневой деривации имеет значение

также для компьютерной лингвистики и вопросов усвоения языка). Кроме того, подобными вопросами задаются в настоящее время и специалисты по когнитивной науке: ср. [Gallistel (в печати)] о противопоставлении «репрезентационных» и «вычислительных» (процедурных) моделей мышления: как замечает Б. Вокс (личное сообщение), по сути в когнитивной науке существует противопоставление двух моделей обучения. В одной из них животные (и, вероятно, люди) усваивают некоторое статическое знание о том, что «нечто является верным» (в этом случае никакой многоуровневости не требуется), в другой знания являются процедурными, они «обучаются делать что-то» (а это подразумевает деривационную структуру). В любом случае вопрос о когнитивных основах ТО до сих пор не является до конца решенным.

- В какой мере все же параллельны языковые структуры на разных уровнях? Мы уже упоминали о синтаксических моделях ТО, подразумевающих архитектуру, не идентичную той, что используется в фонологии. Насколько вообще можно описать в рамках этого «единого» аппарата факты разных языков? Эмпирическую проверку ТО до сих пор нельзя считать вполне завершенной.

Наконец, следует сказать несколько слов о тех вопросах, которые в рамках ТО скорее не рассматриваются или к которым исследователи только начинают подходить.

- Постулируемая универсальность набора ограничений мешает вводить ограничения, как-либо связанные с конкретно-языковыми сущностями (например, конкретными морфемами или грамматическими классами): так, например, при анализе данных атлантического языка серер в [McLaughlin 2000] приходится постулировать некое ограничение, действующее исключительно в некотором подмножестве именных классов этого языка; это, конечно, несколько крайний случай, но тем не менее в работах по ТО часто встречаются ограничения, явно или неявно связанные с элементами словаря конкретного языка. Ср. [Green 2005] о границах применимости ТО: в этой работе предлагается серьезно ограничить «компетенцию» ТО (в частности, выведя за ее границы феномены наподобие «непрозрачности») и исключить из нес частноязыковые правила, невыводимые из более общих принципов. С другой стороны, существуют подходы, позволяющие создавать «морфологически индексированные» ограничения на основе достаточно общих шаблонов, например [Pater 2008].
- В частности, у ТО возникают заметные сложности с нетривиальными чередованиями, не описываемыми с помощью сколько-нибудь общих условий на поверхностную форму (как, например, различные морфонологические чередования в славянских языках, ср. [Sanders 2003] о польском).
- В последнее время наблюдается обращение ТО к тем вопросам, которые раньше не рассматривались, в частности к вопросу о природе фонологических репрезентаций. Какого рода фонологическими единицами оперирует теория? Какие «сегменты» находятся в глубинных представлениях (в частности, нужна ли в теории оптимальности неполная спецификация признаковой структуры, игравшая столь важную роль в деривационных подходах, ср. [Vaux 2003]), какова природа поверхностно-фонологического представления и как оно взаимодействует с фонетическим компонентом языка? Начинается работа по исследованию сегментных наборов в языках мира и природы репрезентации, необходимой для наиболее адекватного их объяснения (ср. [Morén 2007]), а это значит, что вновь поднимается, пусть и в другом виде, «проблема дубликатов»: какие наблюдаемые феномены связаны с процессом деривации, а какие – с природой исходных представлений?

В этом кратком обзоре, к сожалению, невозможно дать представление обо всех аспектах современной ТО, поэтому в заключение мы приведем небольшой список литературы, позволяющей получить представление о ТО и снабженной более или менее подробной библиографией. Очень многие работы по ТО собраны в Rutgers Optimality Ar-

chive [ROA] (<http://roa.rutgers.edu>); в дальнейшем те работы, которые там можно найти, снабжены ссылкой на номер в этом архиве.

- Основные положения ТО подробно изложены в [Prince, Smolensky 1993/2002; McCarthy, Prince 1993a]; важную роль в развитии теории сыграли также работы [McCarthy, Prince 1993b; 1994]; многие важные работы удобно собраны в хрестоматии [McCarthy 2004b].
- Среди «вторичных» источников следует отметить сборник [Archangeli, Langendoen 1997] (содержит главы о различных языковых уровнях), подробный учебник [Kager 1999] и монографический обзор [McCarthy 2001].
- Вопросу о ТО и усвоения языка посвящена монография [Tesar, Smolensky 2000]; вероятностные методы вводятся в известной статье [Boersma, Hayes 1999].
- Хороший обзор проблематики ТО в синтаксисе дан в книге [Legendre et al. 2001]; ср. также [Sells 2001] и важную статью [Grimshaw 1997].

СПИСОК ЛИТЕРАТУРЫ

- Кодзасов, Кривнова 1982 – С.В. Кодзасов, О.Ф. Кривнова. Современная американская фонология. М., 1982.
- Князев 1999 – С.В. Князев. О критериях слогоделения в современном русском языке: теория волны сонорности и теория оптимальности // ВЯ. 1999. № 1.
- Akinlabi 1996 – A. Akinlabi. Featural affixation // Journal of linguistics. V. 31. 1996.
- Alderete 1998 – J. Alderete. Morphologically governed accent in optimality theory. PhD dissertation. University of Massachusetts. Amherst, 1998.
- Anderson 1985 – S. Anderson. Phonology in the twentieth century. Theories of rules and theories of representations. Chicago, 1985.
- Archangeli, Langendoen 1997 – D. Archangeli, D.T. Langendoen. Optimality theory: An overview. Oxford, 1997.
- Beckman 1998 – J. Beckman. Positional faithfulness. PhD dissertation. University of Massachusetts. Amherst, 1998. ([ROA–234]. Отрывки перепечатаны в [McCarthy 2004b]; полный текст – ROA–234).
- Bermúdez-Otero (в печати) – R. Bermúdez-Otero. Stratal optimality theory. Oxford (в печати).
- Blumenfeld 2006 – L. Blumenfeld. Constraints on phonological interactions. PhD dissertation. Stanford University, 2006. [ROA–877].
- Boersma, Hayes 2001 – P. Boersma, B. Hayes. Empirical tests of the gradual learning algorithm // Linguistic inquiry. V. 32. 2001. [ROA–348].
- Broekhuis, Vogel 2006 – H. Broekhuis, R. Vogel (eds.). Optimality theory and minimalism: A possible convergence? // Linguistics in Potsdam. V. 26. Potsdam, 2006 [ROA–892].
- Chomsky, Halle 1968 – N. Chomsky, M. Halle. The sound pattern of English. New York, 1968.
- Gallistel (в печати) – R. Gallistel. The nature of learning and the functional architecture of the brain // Q. Jing et al. (eds.). Psychological science around the world. V. 1. Proceedings of the 28th international congress of psychology. Sussex (в печати).
- Grimshaw 1997 – J. Grimshaw. Projection, heads, and optimality // Linguistic inquiry. V. 28. 1997. [ROA–68].
- Green 2005 – A. Green. Phonology limited (рукопись). [ROA–745].
- Hammond 1997 – M. Hammond. Optimality theory and prosody // D. Archangeli, D. Langendoen. Theory: An overview. Oxford, 1997.
- Haspelmath (в печати) – M. Haspelmath. Parametric versus functional explanations for syntactic universals // T. Biberauer, A. Holmberg (eds.). The limits of syntactic variation. Amsterdam (в печати).
- Hooper 1976 – J.B. Hooper. An introduction to natural generative phonology. New York, 1976.
- Kager 1999 – R. Kager. Optimality theory. Cambridge, 1999.
- Karttunen 2006 – L. Karttunen. The insufficiency of paper-and-pencil linguistics: the case of finnish prosody // R.M. Kaplan, M. Butt, M. Dalrymple, T. Holloway King (eds.). Intelligent linguistic architectures: Variations on themes. Stanford, 2006. [ROA–818].
- Keller, Asudeh 20002 – F. Keller, A. Asudeh. Probabilistic learning algorithms and optimality theory // Linguistic inquiry. V. 33. 2002. [ROA–675].

- Kenstowicz, Kisseberth 1977 – *M. Kenstowicz, Ch. Kisseberth*. Topics in phonological theory. New York, 1977.
- Kenstowicz, Kisseberth 1979 – *M. Kenstowicz, Ch. Kisseberth*. Generative phonology: Description and theory. New York, 1979.
- Kirchner 1996 – *R. Kirchner*. Synchronic chain shifts in optimality theory // *Linguistic inquiry*. V. 27. 1996.
- Lass 1984 – *R. Lass*. Phonology: an introduction to basic concepts. Cambridge, 1984.
- Legendre et al. 2001 – *G. Legendre, J. Grimshaw, S. Vikner* (eds.). Optimality-theoretic syntax. Cambridge, 2001.
- Martinet 1956 – *A. Martinet*. Economie des changements phonetiques. Bern, 1956.
- McCarthy 1998 – *J. McCarthy*. Sympathy and phonological opacity. Amherst, 1998.
- McCarthy 2001 – *J. McCarthy*. A thematic guide to optimality theory. Cambridge, 2001.
- McCarthy 2003 – *J. McCarthy*. Comparative markedness // *Theoretical linguistics*. V. 29. 2003.
- McCarthy 2004a – *J. McCarthy*. The length of stem-final vowels in colloquial Arabic // *M. Alhawary, E. Benmamoun*. Perspectives on Arabic linguistics XVII–XVIII. Amsterdam, 2004. [ROA–616].
- McCarthy 2004b – *J. McCarthy* (ed.). Optimality theory in phonology: A reader. Oxford, 2004.
- McCarthy 2007 – *J. McCarthy*. Hidden generalizations: Phonological opacity in optimality theory. London, 2007.
- McCarthy, Prince 1993a – *J. McCarthy, A. Prince*. Prosodic morphology: Constraint interaction and satisfaction. Rutgers University center for cognitive science technical report. 3. 1993. [ROA–482].
- McCarthy, Prince 1993b – *J. McCarthy, A. Prince*. Generalized alignment. Rutgers University center for cognitive science technical report. 7. 1993. [ROA–7].
- McCarthy, Prince 1994 – *J. McCarthy, A. Prince*. The emergence of the unmarked: Optimality in prosodic morphology // *M. González* (ed.). Proceedings of the North East // *Linguistics society*. V. 24. Amherst, 1994. [ROA–13].
- McCarthy, Prince 1995 – *J. McCarthy, A. Prince*. Faithfulness and reduplicative identity // *J. Beckman, S. Urbanczyk, L.W. Dickey* (eds.). University of Massachusetts occasional papers in linguistics 18: Papers in optimality theory. Amherst, 1995. [ROA–60].
- McLaughlin 2000 – *F. McLaughlin*. Consonant mutation in reduplication in Seereer-Siin // *Phonology*. V. 17. 2000.
- Morén 1997 – *B. Morén*. The division of labor between segment-internal structure and violable constraints // *S. Blaho, P. Bye, M. Krämer* (eds.). Freedom of analysis? Berlin, 1997. [ROA–845].
- Moreton 2004 – *E. Moreton*. Non-computable functions in optimality theory // *J. McCarthy*. Optimality theory in phonology: A reader. Oxford, 2004. [ROA–364].
- van Oostendorp 2007 – *M. van Oostendorp*. Derived environment effects and consistency of exponence // *S. Blaho, P. Bye, M. Kramer* (eds.). Freedom of analysis? Berlin, 2007. [ROA–853].
- Orgun 1996 – *C.O. Orgun*. Correspondence and identity constraints in two-level optimality theory // *Proceedings of WCCFL*. 14. Stanford, 1996.
- Pater 2008 – *J. Pater*. Morpheme-specific phonology: constraint indexation and inconsistency resolution // *S. Parker* (ed.). Phonological argumentation: Essays on evidence and motivation. London, 2008.
- Prince, Smolensky 1993 / 2002 – *A. Prince, P. Smolensky*. Optimality theory: Constraint interaction in generative grammar. Rutgers University, The Johns Hopkins University (Ms.), 2002. [ROA–537].
- Sanders 2003 – *N. Sanders*. Opacity and sound change in the Polish lexicon. PhD dissertation. University of California, Santa Cruz, 2003. [ROA–603].
- Sells 2001 – *P. Sells* (ed.). Formal and empirical issues in optimality theoretic syntax. Stanford, 2001.
- Sells 2003 – *P. Sells*. The INPUT and faithfulness in OT syntax // *J. Spenader, A. Eriksson, Ö. Dahl* (eds.). Proceedings of the workshop on variation within optimality theory. Stockholm, 2003.
- Speas 1997 – *M. Speas*. Null pronouns and control // *D. Archangeli, D. Langendoen*. Theory: An overview. Oxford, 1997.
- Steriade 1988 – *D. Steriade*. Reduplication and syllable transfer in Sanskrit and elsewhere // *Phonology*. V. 5. 1988.
- Tesar, Smolensky 2000 – *B. Tesar, P. Smolensky*. Learnability in optimality theory. Cambridge (Mass.), 2000.
- Vaux 2003 – *B. Vaux*. Syllabification in Armenian. Universal grammar, and the lexicon // *Linguistic inquiry*. V. 34. 2003.