Available online at www.sciencedirect.com

ScienceDirect

Lingua 142 (2014) 66–75

Lingua

www.elsevier.com/locate/lingua

The origin of OT constraints

Jeroen van de Weijer

Shanghai International Studies University, 550 Dalian Road (W), Shanghai 200083, China

Received 22 April 2013; received in revised form 15 November 2013
Available online 22 February 2014

Abstract

In this paper I argue that the two main families of phonological constraints in Optimality Theory, markedness and faithfulness constraints can be derived from more basic cognitive principles operative in language acquisition and lexical storage. In such an approach, the role of the concepts of performance and competence are integrated: the second is derived from the former, representing a step on the way to a truly non-derivational theory of grammar based on language use.

Keywords: Phonology; Optimality Theory; Constraints; Language use; Acquisition; Lexicon

1. Introduction

The question of whether the grammar of languages is wholly or partially innate, or, more specifically, which units, notions or concepts are innate, if any, is one of the most vehemently debated issues in the history of linguistics. In the current paper, I argue that key parts of what is currently the most successful theory in phonology, Optimality Theory (OT) (Prince and Smolensky, 1993 [2004]), need not be assumed to be innate (as has usually been done), but can be derived from more general principles operative in language acquisition and lexical storage. This is not intended as an argument against this theory: on the contrary, it is intended to strengthen the basis of OT by uncovering its relations with other theories of language and cognition, in particular usage-based theories, as well as making a closer connection between OT and findings in psycholinguistics (as suggested by Jackendoff (2007); see also Goldinger (2007), Hawkins (2003), McLennan et al. (2003) for approaches to a more unified theory of language). Specifically, I will argue that markedness and faithfulness constraints, as well as some other key assumptions in Optimality Theory need not be stipulated as innate, but fall out quite naturally from the language acquisition process and the communicative purposes for which language is used. Specifically, constraints turn out to be generalizations across the data to which the language-learning is exposed. Two things are crucial: word frequency in lexical storage plays a role (especially for markedness constraints), so that words that are more frequent (especially in data to which the child is exposed early) and storage of similar (though not identical) forms plays a role (especially for faithfulness constraints). Both ''ingredients'' are present in Exemplar Theory (Johnson (1997), Pierrehumbert (2001), Bybee (2006) and references cited there; see below), so we will propose to combine the Exemplar view of storage with the Optimality concept of grammar. Other approaches to derive markedness constraints certainly exist, though they are not based on lexical storage, such as those pursued by constraint induction models (e.g. Hayes (1996), Hayes and Wilson (2008), Adriaans and Kager (2010), and stochastic Optimality Theory (Boersma (1997: Ch. 14) and subsequent work)). A proposal to derive correspondence constraints from the perceptibility of different contrasts is made by Steriade (2009). The difference between these proposals and the present one is that the

E-mail address: jeroen.van.de.weijer@gmail.com.

0024-3841/$ – see front matter
http://dx.doi.org/10.1016/j.lingua.2014.01.007

present one tries to relate all constraints (markedness and faithfulness) to the Exemplar model of lexical storage. In this way, constraints need not be assumed to be innate.

This paper is organized as follows: in Section 2 I discuss the nature of grammar, with particular attention for aspects of universality and innateness, the rule (or constraint) system and the role of the lexicon. In Section 3, I discuss how markedness constraints could be derived in the course of first language acquisition. In Section 4 I show how faithfulness constraints, the other main class of Optimality constraints, could emerge if we adopt an alternative, better-articulated view of lexical storage. The relevance of this alternative view, couched in Exemplar Theory, is outlined in Section 5. In Section 6 I tie Exemplar storage and OT constraints together and explore the implications for our conception of the overall architecture of grammar. Section 7 briefly concludes.

2. Grammar

Perhaps the most fundamental part of a phonological grammar in generative linguistics is a mechanism of converting inputs into outputs. Let us briefly review its most fundamental aspects, because they will be relevant later on. The output level is sometimes also referred to as the ''phonetic'' or ''surface'' level and the input is ''phonological'', ''underlying'' or ''lexical''—these terms are used in slightly different ways by different authors. Sometimes intermediate levels are assumed, as in Lexical Phonology (Kiparsky, 1982) and some versions of Optimality Theory (Rubach, 2000, 2003). The assumption of at least two levels forms a core idea of this branch of linguistics, and is present in both earlier rule-based approaches, which we will refer to as ''standard generative phonology'' (Chomsky and Halle (1968)), as well as in constraint-based theories such as OT (Prince and Smolensky, 1993 [2004]). Thus, both standard generative phonology and OT are generative theories, in the sense that an output is generated on the basis of an input, the difference being that in the former generation is achieved by rule application (derivation) and in the latter by selection of the most successful (''optimal'') candidate from a pool of contenders. OT can thus be considered a non-derivational, generative theory of grammar.

What do speakers know if they know the grammar of their language and what is the balance between innate parts of this knowledge and parts acquired in the course of language acquisition? I will examine this question for both rule-based generative phonology and for constraint-based generative phonology, to highlight the fact that both types of theories are basically equivalent in their assumptions about such notions, and second, that constraint-based approaches nevertheless have a number of important advantages over earlier rule-based theory.

In rule-based generative phonology, an input is converted into an output by way of application of one or more phonological rules. The input is the phonological form, e.g. Dutch /hɔnd/ 'dog', to which one or more phonological rules may apply, e.g. Final Devoicing [–son] → [–voice] / ___] (''obstruents are devoiced at the end of the syllable''), which turns phonological /hɔnd/ into phonetic [hɔnt] (phonological forms appear between forward slashes, phonetic forms in brackets). The phonological rule does not apply to the underlying form of the plural, /hɔndən/ 'dogs', because here /d/ is not syllable-final. In standard generative phonology, then, knowledge of the grammar corresponds to knowing the underlying forms of the language as well as the phonological rules, both of which have to be learned in the course of language acquisition. In this theory, notions such as ''syllable'' and distinctive features like ''[–voice]'' might still be innate (since they appear in the formulation of the rule, we must ask the question where these notions come from), as well as the overall architecture of the grammar which consists, minimally, of an underlying level and surface forms, and makes certain assumptions about the relation between the phonology, syntax, morphology and semantics. Notice that the surface forms themselves are not part of the grammar, nor do they need to be remembered or learned as part of the lexicon (which contains underlying forms). If output forms were learned, much of the generative mechanism, viz. the underlying forms as well as the rules, would be superfluous. Of course, in generative grammar, the rules are argued to express generalizations that a native speaker knows about his or her language. Evidence for this comes from the mis- or over-application of rules when a child produces *sheeps* as the plural of *sheep*, or the past tense *swimmed*. If such patterns can be expressed in other ways than the familiar underlying forms + rules approach, some of the need for multiple levels falls down.

Constraint-based approaches, which were developed in response to standard generative phonology, are also generative in the sense that they map output to input, with a set of constraints and a selection mechanism taking the place of rules. The status of output forms remains essentially unchanged. As is well known, an output form (e.g. [hɔnt]) is mapped onto an input /hɔnd/ by means of an evaluation mechanism, which in the standard version of Optimality Theory (Prince and Smolensky, 1993 [2004]) consists of a set of universal, ranked constraints. The winning output that best fits the constraint hierarchy for a particular language is selected (from a set of candidate outputs that is infinite in standard OT, (Kager, 1999: 8)) and will be the one that is pronounced. A constraint parallel to the Final Devoicing rule (see above) will be assumed to be high-ranked in such a language, e.g. *Voiced-Coda (''final voiced obstruents are not permitted'' ((Kager, 1999: 14); or a combination of constraints with the same effect); see e.g. Lombardi (1995)). In the constraint-based approach, the native speaker's grammar consists of a lexicon of input forms (the Dutch speaker has to learn that the correct input for 'dog' is /hɔnd/ and not /hɔnt/) and a number of constraints in the correct, language-specific hierarchy. In