

## Violable Constraints in Language Production: Testing the Transitivity Assumption of Optimality Theory

Daniel J. Guest, Gary S. Dell, and Jennifer S. Cole

*University of Illinois at Urbana-Champaign*

Optimality Theory (OT) (Prince & Smolensky, 1993) characterizes linguistic knowledge as a ranked set of constraints that select the best possible output form of a word given a particular input. OT assumes that constraints are ordered transitively with respect to their violability. An artificial language learning paradigm was used to test this assumption by teaching participants to pronounce words that provided evidence about three constraints affecting the stress patterns of words. The words demonstrated that the first constraint outranked the second and the second outranked the third. The relationship between the first and third could only be derived from the transitive nature of the system. Three experiments tested whether speakers could determine the stress patterns of words requiring knowledge of the relationship between these two constraints. Evidence was found for a transitively ordered constraint system as well as a system that stores commonly heard stress patterns as metrical templates. © 2000 Academic Press

*Key Words:* Optimality Theory; stress; phonology; language acquisition; language production; transitive inference.

Optimality Theory (Prince & Smolensky, 1993), or OT, is an important new paradigm in linguistics. In phonology OT has quickly become the dominant framework. Since the early 1990s, numerous conference talks and journal articles in this area have presented their ideas in OT terms. OT has also proven to be a useful framework for syntactic (e.g., Pesetsky, 1997; Speas, 1997) and morphological (Russell, 1997) analyses.

The OT approach diverges from traditional linguistic analyses by proposing that grammars consist of violable constraints rather than rules.

Portions of this work were presented at the 2nd annual meeting of the Mid-Continental Workshop on Phonology, the 69th Annual Meeting of the Midwestern Psychological Association, and the 38th Annual Meeting of the Psychonomics Society. The research was supported by National Science Foundation Grants SBR 93-19368 and SBR 98-73450 and National Institutes of Health Grants DC-00191 and HD-21011. The authors thank Kay Bock, Cindy Fisher, Stefan Frisch, Susan Garnsey, and Paul Smolensky for helpful discussions and Arva Bensaheb for assistance collecting data.

Correspondence and reprint requests should be sent to Daniel Guest at the Beckman Institute, University of Illinois at Urbana-Champaign, 405 North Mathews Avenue, Urbana, IL 61801. E-mail: [dguest@s.psych.uiuc.edu](mailto:dguest@s.psych.uiuc.edu).

In OT, Universal Grammar defines a set of universal constraints that function in all languages to impose conditions on the structure of linguistic surface forms. In phonology, the constraints govern the features and structures that determine a language's sound forms. For example, the Weight to Stress Principle, or HEAVY constraint, requires that heavy consonant–vowel–consonant (CVC) syllables be stressed. A defining characteristic of OT is that the universal constraints often impose mutually incompatible requirements on an individual surface form. For example, consider the conflict that can arise in a language with heavy CVC syllables, and in which stress is located on the initial syllable of a word, due to a constraint termed ALIGN-FOOT(left) that requires a stress foot at the left edge of the word. If such a language includes a word with a heavy CVC syllable in second position, the combined effects of ALIGN-FOOT(left) and the HEAVY constraint would be to place stress on both the initial and second syllables. This pattern of adjacent stresses violates another constraint operative in stress systems, CLASH, which bans any structure that contains adjacent stressed syllables.

Thus, there is a conflict among HEAVY, ALIGN-FOOT(left), and CLASH: These three constraints cannot be simultaneously satisfied in a word with a heavy CVC second syllable. OT resolves constraint conflict by ranking all constraints in a transitive dominance hierarchy; in the case of conflict, priority is given to the higher ranked constraint, at the expense of a violation of the lower ranked constraint. In the example just mentioned, ranking CLASH and HEAVY higher in the constraint hierarchy than the constraint requiring stress on the initial syllable, ALIGN-FOOT(left), will result in a surface form with stress on the second but not the first syllable of this word. The relative ranking of constraints in the hierarchy varies across languages, which accounts for variation in the surface phonological patterns found in different languages.

In this paper, we present three experiments that tested a key assumption of OT, namely, that the constraint ranking relation is transitive: If Constraint A outranks Constraint B, and B outranks C, then A also outranks C. In the experiments, participants learned constraints about how to stress words in artificial languages derived from OT grammars. They were then tested on novel items requiring the application of transitivity. As will be seen, we found evidence for transitivity and, thereby, for a system that determines pronunciations from ranked general constraints. But we also found evidence for another process in learning words' stress patterns, a process in which novel words are stressed by matching them to already stored stress patterns or templates. Thus, our research—subject to the usual caveats associated with artificial language learning experiments—is consistent with a dual-system approach to linguistic performance, one system associated with a grammar and one with a lexicon.

Our studies asked whether OT principles have utility in explaining how people learn to produce word forms. As such, the research is psycholinguistic in the traditional sense: A linguistic theory is examined for its "psychological reality." OT is particularly promising in this vein. The theory has links to, and has evolved at

least partially from, connectionist theories of cognition (Rumelhart & McClelland, 1986), and these theories have had considerable impact on psycholinguistics (see Christiansen & Chater, in press, for review). OT is very much a part of the *Zeitgeist* in cognitive science that eschews hard-and-fast rules in favor of soft or violable constraints. At the same time, the strict ranking of constraints represents a crystallization of a set of violable constraints into a systematic and learnable language.

Before describing the experiments, we provide an overview of OT and review previous psycholinguistic research on OT and on the psychology of transitive inference.

### OVERVIEW OF OPTIMALITY THEORY

The architecture of OT consists of a set of universal constraints, a relation that ranks constraints in a dominance hierarchy, a mechanism for generating candidate surface output forms that are evaluated against the lexical input and for surface well-formedness by the phonological constraints, and a mechanism for constraint evaluation (see Archangeli & Langendoen, 1997, for an overview of the theory).

The claim that constraints are provided by Universal Grammar means that constraints must in principle be violable; if they were not, then all languages would conform to the universal constraints in a similar fashion and display identical surface phonological patterns. Language-specific constraint ranking is thus a critical aspect of the theory, since it provides the mechanism for determining which of two conflicting constraints will be violated in a given language.

At this juncture it is useful to introduce the notation that is adopted in this paper. Consider again the example from above of conflict between the constraints HEAVY and ALIGN-FOOT(left). We stated that the conflict that arises in the evaluation of a word with a CVC syllable in second position can be resolved by ranking CLASH and HEAVY higher than ALIGN-FOOT(left). Ranking is noted as: Constraint A  $\gg$  Constraint B, read as "Constraint A outranks Constraint B." Evaluation of surface candidates for such a lexical input is illustrated

TABLE 1

OT Table Demonstrating Evaluation of Surface Candidates

	cv.cvc.cv	cṿ.cvc̣.cv	c̣v.c̣vc̣.cv	c̣ṿ.c̣vc̣.cv
CLASH				*!
HEAVY			*!	
ALIGN-FOOT(left)	*			

in Table 1, which presents an OT table in the format proposed by Bernhardt and Stemberger (1998) (this format diverges slightly from that of the mainstream OT literature, but provides for a potentially cleaner presentation. The reader is referred to Bernhardt & Stemberger, 1998, particularly Chapter 4, for the motivations behind the changes).

The lexical input, or underlying, form is in the top, left corner, with candidate surface forms following across the top row. The phonological constraints are listed down the left-most column, in a vertical order that reflects the dominance ranking. An asterisk (\*) in a cell marks constraint violation for the candidate that heads the column. An exclamation point following an asterisk (\*!) indicates a “fatal” constraint violation that renders the candidate suboptimal relative to other candidates that satisfy the constraint evaluated in that, or higher, rows. All cells in rows lower than the fatal violation are shaded, indicating that their marks are not relevant to the determination of the optimal candidate. The winning candidate is the one that fares the best under the highly ranked constraints and is indicated with the pointer symbol (☞).

In the example given in Table 1, “CV” denotes an open consonant–vowel syllable, “CVC” denotes a closed heavy syllable, and “.” denotes a syllable boundary. The second candidate violates HEAVY, incurring a fatal violation on this highly ranked constraint. The third candidate, which stresses both the first and second syllables, violates CLASH, also a fatal violation. The winning candidate, then, is the first candidate, which satisfies HEAVY, despite its violation of the lower ranked ALIGN-FOOT(left). Changing the domination order will

TABLE 2

OT Table Demonstrating the Effects of Reordering Constraints

	cv.cvc.cv	cṿ.cvc̣.cv	☞c̣ṿ.c̣vc̣.cv	c̣ṿ.c̣vc̣.cv
CLASH				*!
ALIGN-FOOT(left)		*!		
HEAVY			*	

change which candidate is chosen as being the most optimal. This is demonstrated in Table 2, where the form c̣ṿ.c̣vc̣.cv is determined to be the optimal candidate, and therefore the designated surface form, under the new ranking CLASH, ALIGN-FOOT(left) >> HEAVY for the input.

### THE ROLE OF TRANSITIVITY IN OT

The ranking relation in OT is transitive, which means that if  $A \gg B$ , and  $B \gg C$ , then  $A \gg C$ . Transitivity is a crucial feature of OT constraint ranking. Without transitivity, evaluation can fail to identify any candidate as optimal for a given lexical input form. Consider the following schematic example. Assume a grammar with three constraints, A, B, and C, and the ranking  $A \gg B$  and  $B \gg C$ . Now assume that the grammar has the nontransitive ranking  $C \gg A$ . Table 3 shows the evaluation of three distinct candidate forms under these three constraints. Note that constraint C is listed first in the column of constraints, although that order does not reflect its ranking beneath constraint B.

This table does not indicate a winner, because each candidate must be rejected on the basis of the proposed ranking. Output-a cannot be the winner because it loses to output-b under the

TABLE 3

Evaluation of Candidates under a Nontransitive Constraint Ranking

Input	Output-a	Output-b	Output-c
Constraint C			*!
Constraint A	*!		
Constraint B		*!	

ranking  $A \gg B$ . Similarly, output-b cannot be the winner because it loses to output-c under the ranking  $B \gg C$ . Finally, output-c cannot be the winner either because it loses to output-a under the ranking  $C \gg A$ . Selecting any one of these candidates as the winner entails contradicting one of the assumed constraint rankings. Therefore, if all three constraint rankings are valid then there is no winner in a case such as this, and the grammar is incomplete because it cannot guarantee a decisive outcome of evaluation for an individual input. This example demonstrates that constraint ranking must be transitive in OT in order for any OT grammar to be complete—a minimum requirement for a formal theory.

Transitivity is also important for the learnability of an OT grammar. To learn a grammar, one must learn how each constraint ranks with respect to all other constraints. In a system with  $n$  constraints, there are  $n!$  possible constraint rankings which, on the surface, makes for a daunting learning task. However, because of transitivity, at a minimum, only  $n(n-1)$  informative constraint interactions are needed to learn the correct ranking for the system (Tesar & Smolensky, 1993; Prince & Smolensky, 1997).

#### PSYCHOLINGUISTICS AND OT

Although a large amount of work has been done in phonology utilizing Optimality Theory (at the time this article was written, there were 319 papers listed in the Rutgers Optimality Archive alone), there has been relatively little psycholinguistic research on OT. Bernhardt and Stemberger (1998), one of the few psycholinguistic works employing OT, applied the framework to a variety of phenomena in phonological and prosodic development. In addition, they were able to employ constraint rankings to describe the deficits of children with phonological delays. This is significant because they did not need to posit ad-hoc constraints to handle the language disorders of these children. Instead, language performance was explained by finding the appropriate aberrant constraint ranking for each child and seeing how these rankings might change through successful intervention. Other work on phonological development has also

emphasized the use of violable constraints (Lev-elt, 1994).

OT has also been used to interpret psycholinguistic investigations of the role of the syllable in speech perception (Hammond's, 1995, interpretation of Cutler, Mehler, Norris, & Segui, 1983, 1986) and in investigations of the structure of Hebrew words (Berent, Everett, & Shimron, 1998).

#### TRANSITIVE INFERENCE

Transitive inference, in various forms, appears to be a pervasive information processing ability. The essence of such an inference is as follows: The truth of a relation between two items,  $A > C$ , can be inferred from the truth of  $A > B$  and  $B > C$ . Research has demonstrated that transitive inference occurs in various animal species (e.g., pigeons, Steirn, Weaver, & Zentall, 1995; Weaver, Steirn, & Zentall, 1997; macaque monkeys, Treichler & Tilburg, 1996; squirrel monkeys, McGonigle & Chalmers, 1977, 1992; chimpanzees, Boysen, Berntson, Shreyer, & Quigly, 1993; and rats, Davis, 1992) and in children (e.g., 7 year olds, Piaget, 1967; 4 year olds, Trabasso, 1977; and 22 month olds, Lipkens, Hayes, & Hayes, 1993). In addition, research has investigated mechanisms for these inferences in particular domains. For example, studies have examined whether transitive inferences about ordered sets involve the manipulation of propositions (Clark, 1969) or images (Huttenlocher, 1968).

There have also been suggestions that transitive inferences can be made implicitly (e.g., Lewicki, Hill, & Czyzewska, 1994). Mediated priming (lion primes tiger, and tiger primes stripes, therefore lion primes stripes; e.g., Balota & Lorch, 1986; McNamara, 1992; McKoon & Ratcliff, 1992; O'Seaghdha & Marin, 1997; Shelton & Martin, 1992) can be thought of as a kind of implicit transitive inference, where the relevant relationship is one of association. In a similar vein, people's ability to verify statements such as "canaries are animals" can also be interpreted as implicit transitive inference, provided that the mechanism for such verification are the previously stored proposi-

tions that “canaries are birds” and “birds are animals” (Collins & Quillian, 1969).

For our purposes, the most important aspects of this literature are the data suggesting that transitive inference occurs in young children and may occur implicitly. Phonological development largely occurs between birth (or before) and age 4 (Bernhart & Stemberger, 1998; Oller, 1980; Stark, 1980; Vihman, 1996) and the acquisition process and the resulting knowledge are generally considered implicit. If speakers compute the phonological forms of words through a system resembling an OT grammar, one might expect to see evidence for transitive inference as the system is acquired. Consequently, our studies investigated how people learn the form of words.

We used a laboratory learning paradigm in which adult speakers learned the stress patterns of syllable strings from an artificial language. The stress patterns were derived from an OT grammar, and the training illustrated the ranking of particular pairs of constraints, in particular that Constraint A  $\gg$  B and that B  $\gg$  C. (The particular constraints varied in different language conditions). The participants were then asked to pronounce novel syllable strings that put A and C into conflict for the first time. If speakers produce the stress pattern predicted by the grammar for these test items, they may have done so by using transitivity.

## EXPERIMENT 1

### *Method*

#### *Participants*

Twenty native English-speaking undergraduates from the University of Illinois participated in exchange for credit toward a course requirement. Equal numbers of participants were randomly assigned to one of two language conditions, language A (LA) and language B (LB).

#### *Stimulus Materials*

Two languages were created that demonstrated the interaction among constraints that govern the location of stress in polysyllabic words. Words in each of the two languages

were built up from a lexicon of the same eight syllables, consisting of seven CV syllables and one CVC syllable. Stress placement was governed in each language by a small number of constraints adopted from the linguistic theory of metrical stress. The stress patterns of the two languages were mirror images of one other, and their respective constraint systems were composed of identical, but mirror-image constraints. This opposing nature of the languages enabled them to serve as controls for each other. Participants were trained on words from one of the languages and then were given a common set of test words. Differences on the test items as a function of training can therefore be used to index learning. Importantly, any such differences cannot be ascribed to participants' knowledge of English phonology since participants would be expected to produce similar stress patterns given the same test words if their English phonology was at work.

*Language A stress constraints.* The stress patterns of LA words can be described in terms of four basic properties:

1. The penultimate syllable of the word is stressed.
2. Stress is placed on every odd-numbered syllable counting from the beginning of the word, excluding the final syllable.
3. Every heavy syllable is stressed.
4. Adjacent syllables may not both be stressed (no stress clash).

Each property of LA stress is individually attested in a number of natural language stress systems and can be attributed to one or more constraints that govern the location of stress.<sup>1</sup> Properties 1 and 2 are attributed here to the constraints PENULT (the penultimate syllable of a word must be stressed) and LEFT-ALTERNATING (there must be an alternating pattern of stressed–unstressed stated below).

Both PENULT and LEFT-ALTERNATING

<sup>1</sup> The constraints are described in nontechnical language here, to facilitate ready understanding of the essential stress characteristics of the two experimental languages. All of the constraints described here are modeled after actual constraints proposed for the analysis of metrical stress systems in OT by McCarthy and Prince (1993), drawing heavily on Hayes' (1995) theory of metrical stress. A more detailed characterization of the metrical constraints relevant to these data is found in the Appendix.

are composite constraints, built up from a small set of atomic OT constraints whose interaction governs the location and internal composition of the metrical foot, the structural unit that underlies stress marking in systems with rhythmic stress. The atomic constraints are given in the Appendix, along with an illustration of their individual functions in determining stress placement in LA. Leaving aside heavy syllables for a moment, in words with an even number of syllables, the pattern of alternating stresses required by LEFT-ALTERNATING will locate a stress on the penultimate syllable, as in  $\acute{\sigma}\sigma\acute{\sigma}$  and  $\acute{\sigma}\sigma\acute{\sigma}\sigma$ , where  $\sigma$  refers to a single CV syllable. A form containing an even number of syllables with this stress pattern would satisfy both the PENULT and LEFT-ALTERNATING constraints. For words with an odd number of syllables, there is not, strictly speaking, any stress pattern that could satisfy both PENULT and LEFT-ALTERNATING. Locating a stress on the penultimate syllable and on the odd-numbered syllable that precedes it would result in a stress clash, as in  $\acute{\sigma}\sigma\acute{\sigma}\sigma$ . Stress clash is not tolerated in many natural language stress systems; there are no words in the LA lexicon that present stress on two adjacent syllables. The avoidance of stress clash in LA is attributed to the CLASH constraint (adjacent syllables may not both be stressed).

In LA, CLASH is avoided by locating stress on the penultimate syllable, but not on the odd-numbered syllable that precedes it, as in  $\acute{\sigma}\sigma\sigma\acute{\sigma}$  or  $\acute{\sigma}\sigma\acute{\sigma}\sigma\sigma$ . Words with an odd number of syllables that display this stress pattern satisfy both CLASH and PENULT, but at the expense of a violation of LEFT-ALTERNATING. The interaction between these three constraints can be modeled in OT by ranking CLASH and

TABLE 4

CLASH, PENULT  $\gg$  LEFT-ALTERNATING

	$\sigma\sigma\sigma\sigma$	$\acute{\sigma}\sigma\sigma\acute{\sigma}$	$\acute{\sigma}\sigma\acute{\sigma}\sigma$	$\acute{\sigma}\sigma\sigma\sigma$
CLASH			*!	
PENULT				*!
LEFT-ALTERNATING	*		*	

TABLE 5

CLASH, HEAVY  $\gg$  PENULT

	$\sigma\sigma\sigma_H$	$\acute{\sigma}\sigma\sigma_H$	$\sigma\sigma\sigma_H$	$\sigma\sigma\sigma_H$
CLASH			*!	
HEAVY				*!
PENULT	*			
LEFT-ALTERNATING			*	*

PENULT above LEFT-ALTERNATING, as illustrated in Table 4.

The three constraints discussed so far are by themselves responsible for the stress pattern of all words that contain only light syllables. Heavy syllables introduce a further complication because they attract stress. That is, a heavy syllable is stressed even when it occurs outside of a position that is designated for stress by PENULT and LEFT-ALTERNATING. The stress-attracting behavior of heavy syllables is illustrated by LA words with forms such as  $\acute{\sigma}\sigma\sigma_H$  and  $\acute{\sigma}\sigma\sigma_H\sigma$ , where  $\sigma_H$  refers to a CVC syllable, and is attributed here to the HEAVY constraint (a CVC syllable must be stressed).

The principle of stressing heavy syllables overrides the placement of stress on the penultimate syllable in LA, as examples of the form  $\acute{\sigma}\sigma\sigma_H$  or  $\acute{\sigma}\sigma\sigma_H\sigma$  demonstrate. This result is achieved by ranking the HEAVY constraint over PENULT. These examples also demonstrate the principle of stress clash avoidance, which rules out the possibility of stressing the heavy syllable *in addition to* the penultimate syllable in these forms. The necessary constraint rankings are illustrated in the evaluation of the example from  $\acute{\sigma}\sigma\sigma_H$  in Table 5. The LEFT-ALTERNATING constraint is included in Table 5 for completeness, although because of its low ranking relative to all the other constraints, it plays no active role in this evaluation scheme.

The final property of LA to be discussed concerns exceptions to the principle that requires stress on every heavy syllable. If a word contains two adjacent heavy syllables, stress will fall on only one of them, due to the overriding effect of stress clash avoidance. Examples are words of the form  $\acute{\sigma}_H\sigma_H\sigma\sigma$  and

$\acute{\sigma}_H\acute{\sigma}_H\sigma\sigma$ . The choice of which of the two adjacent heavy syllables to stress is determined by the other ranked constraints. These stress patterns are accounted for by ranking CLASH over HEAVY, as is demonstrated with the form  $\acute{\sigma}_H\sigma_H\sigma\sigma$  in Table 6.

Note that in Table 6, the third candidate,  $\sigma_H\acute{\sigma}_H\sigma\sigma$ , is eliminated in favor of the winning candidate on the basis of its two violations of LEFT-ALTERNATING. While neither the first or third candidates succeed in placing a stress on the third syllable, the third candidate also suffers from its failure to stress the first syllable.

*Language B stress constraints.* The words in Language B (LB) exhibit stress patterns that are the mirror images of the LA stress patterns. Thus, the four basic properties of LB stress patterns are

1. The second (pen-initial) syllable of the word is stressed.
2. Stress is placed on every odd-numbered syllable counting from the end of the word, excluding the initial syllable.
3. Every heavy syllable is stressed.
4. Adjacent syllables may not both be stressed (no stress clash).

As with LA, each property of LB stress is individually attested in natural language stress systems and can be attributed to one or more constraints that govern the location of stress. This system can be accounted for with a variant of the constraint set for LA. In particular, the PENULT and LEFT-ALTERNATING constraints that govern LA have their counterparts in the constraints PEN-INITIAL (the pen-initial syllable of a word must be stressed) and RIGHT-ALTERNATING (there must be an alternating pattern of stressed–unstressed syllables

TABLE 7

CLASH, PEN-INITIAL  $\gg$  RIGHT-ALTERNATING

	$\sigma\sigma\sigma\sigma$	$\acute{\sigma}\sigma\sigma\sigma$	$\sigma\sigma\sigma\acute{\sigma}$	$\sigma\sigma\sigma\acute{\sigma}$
CLASH			*!	
PEN-INITIAL				*!
RIGHT-ALTERNATING	*		*	

bles starting the right edge of the word going leftward, excluding the initial syllable) that govern LB stress. The CLASH and HEAVY constraints play the same role in both experimental languages.

The constraint interactions in LB are the same as those in LA, substituting PEN-INITIAL and RIGHT-ALTERNATING for PENULT and LEFT-ALTERNATING, respectively. For example, in LB the ranking of CLASH and PEN-INITIAL over RIGHT-ALTERNATING accounts for the stress pattern on words with an odd number of light syllables, as in  $\sigma\sigma\sigma\acute{\sigma}$  or  $\sigma\sigma\sigma\acute{\sigma}\sigma$ , as shown in Table 7.

As can be seen by comparing Tables 4 and 7, the substitution of PENULT and LEFT-ALTERNATING in LA with PEN-INITIAL and RIGHT-ALTERNATING in LB produces mirror-image stress patterns in the two languages with identical constraint violations of their respective constraints. This aspect of the two languages can be further examined in Table 8, which lists all of the LA and LB forms that were used to train the participants.

*Words.* The words for the two experimental languages were created from the open (CV) syllables of the musical scale: do, re, mi, fa, so, la, it, do. To this set we added a single heavy (CVC) syllable: “ton,” pronounced [to:n].<sup>2</sup> The

TABLE 6

CLASH  $\gg$  HEAVY

	$\sigma_H\sigma_H\sigma\sigma$	$\acute{\sigma}_H\sigma_H\sigma\sigma$	$\acute{\sigma}_H\acute{\sigma}_H\sigma\sigma$	$\sigma_H\acute{\sigma}_H\sigma\sigma$
CLASH			*!	
HEAVY	*			*
PENULT				
LEFT-ALTERNATING	*		*	**!

<sup>2</sup> When stressed, the open syllables are produced with a tense diphthongal rime, e.g., [dow], [rej]. In unstressed position, subjects varied in their productions, with some tokens produced with reduced vowels, e.g., [də], [rə], and other tokens produced with a tense vocalic rime, e.g., [do], [re], shorter in duration than the stressed counterparts. Thus, the open syllables of the experimental languages are most accurately characterized as CV: (or consonant–vowel–glide), at least in stressed tokens. The open syllables contrast with the heavy syllable primarily in the absence of a (nonvocalic) coda consonant, which is the basis for our

TABLE 8

LA and LB Training Types (Number of Tokens)

Language A			Language B		
CLASH >> HEAVY	HEAVY >> PENULT	PENULT >> LEFT- ALTERNATING	CLASH >> HEAVY	HEAVY >> PEN-INITIAL	PEN-INITIAL >> RIGHT-ALTERNATING
$\acute{\sigma}_H\sigma_H\acute{\sigma}\sigma$ (28)	$\acute{\sigma}_H\sigma\sigma$ (28)	$\acute{\sigma}\sigma\acute{\sigma}\sigma$ (28)	$\acute{\sigma}\sigma\sigma_H\acute{\sigma}_H$ (28)	$\sigma\sigma\acute{\sigma}_H$ (28)	$\acute{\sigma}\sigma\sigma\acute{\sigma}$ (28)
$\acute{\sigma}_H\sigma_H\acute{\sigma}_H\acute{\sigma}\sigma$ (28)	$\acute{\sigma}\sigma\acute{\sigma}_H$ (28)	$\acute{\sigma}\sigma\sigma\sigma\acute{\sigma}\sigma$ (28)	$\acute{\sigma}\sigma\acute{\sigma}_H\sigma_H\acute{\sigma}$ (28)	$\acute{\sigma}_H\sigma\acute{\sigma}$ (28)	$\acute{\sigma}\sigma\sigma\sigma\acute{\sigma}\sigma$ (28)
	$\acute{\sigma}\sigma\acute{\sigma}_H\sigma\sigma$ (28)	$\acute{\sigma}_H\sigma\sigma\sigma\sigma\acute{\sigma}$ (6)		$\sigma\sigma\acute{\sigma}_H\sigma\acute{\sigma}$ (28)	$\acute{\sigma}\sigma\sigma\sigma\sigma\acute{\sigma}_H$ (6)
	$\acute{\sigma}\sigma\sigma\sigma\acute{\sigma}_H\sigma\sigma$ (6)	$\acute{\sigma}\sigma\acute{\sigma}_H\sigma\sigma\sigma\acute{\sigma}$ (28)		$\sigma\sigma\acute{\sigma}_H\sigma\sigma\sigma\acute{\sigma}$ (6)	$\acute{\sigma}\sigma\sigma\sigma\sigma_H\sigma\acute{\sigma}$ (28)
	$\acute{\sigma}\sigma\sigma\sigma\sigma\acute{\sigma}_H$ (28)			$\acute{\sigma}_H\sigma\sigma\sigma\sigma\acute{\sigma}$ (28)	

syllables of the musical scale were chosen because of their familiarity. The sequential order of the syllables was another important feature that was exploited in constructing the words of the experimental languages.

In both LA and LB, the CV syllables within a word were kept in the sequence defined by the musical scale. This sequence defines a loop, where “ti” can be followed by “do,” “do re,” and so on, continuing through the series. Words were created by varying word length and the starting syllable of the word. Words had a minimum of three syllables and a maximum of seven syllables. The CVC syllable “ton” could be placed anywhere in the order. Example words (with no stress indication) are *re mi fa so la*, *la ton ti do re*, and *ton ton mi fa so*.

Words for LA and LB were constructed with the purpose of inducing conflict among the four stress constraints. For example, LA includes words with an odd number of CV syllables, such as “do re mi fa so” or “mi fa so la ti do re,” patterns which were seen above to lead to a conflict between PENULT and LEFT-ALTERNATING. The stress pattern presented on words such as these shows penultimate stress, i.e., “dó re mi fá so” and “mí fa só la ti dó re,” and demonstrates the ranking of CLASH, PENULT >> LEFT-ALTERNATING. Similarly, there were words in LA that demonstrated the ranking of HEAVY >> PENULT, such as “tón

re mi” and “só la tón,” and words that demonstrated the ranking of CLASH >> HEAVY, such as “tón ton mi fá so.”

Two sets of words were created, serving both LA and LB: training words and test words. The training words for the two languages differed only in that the placement of stress and of the CVC syllables occurred in mirror-image positions. Each training word demonstrated a single constraint ranking from a set of three critical rankings: for LA these rankings were CLASH >> HEAVY, HEAVY >> PENULT, and PENULT >> LEFT-ALTERNATING; for LB the demonstrated rankings were CLASH >> HEAVY, HEAVY >> PEN-INITIAL, and PEN-INITIAL >> RIGHT-ALTERNATING. Crucial to the design of the training set was that none of the training words demonstrated a relationship between HEAVY and LEFT-ALTERNATING (LA) or between HEAVY and RIGHT-ALTERNATING (LB).

There were 11 different types of training stimuli for each language, where a type refers to a group of seven words of the same length, with the syllable “ton” in the same position, and with the words differing only in the initial syllable. For example, *do re ton*, *re mi ton*, *mi fa ton*, *fa so ton*, *so la ton*, *la ti ton*, and *ti do ton* are all of the type  $\sigma\sigma\sigma_H$ . Table 8 shows the different types of words that were used in the training set to demonstrate each of the three rankings, for both LA and LB.

The test items probed transitivity in constraint ranking by putting the constraints HEAVY and LEFT-ALTERNATING (RIGHT-

choice of labels: “CV” for the open syllables and “CVC” for the closed syllables. For further discussion of syllable weight, see the Appendix.

TABLE 9

Seven-Syllable Test Item for LA: Conflict between  
HEAVY and LEFT-ALTERNATING

$\sigma\sigma\sigma_{\text{H}}\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma_{\text{H}}\sigma\sigma\sigma$	$\sigma\sigma\sigma_{\text{H}}\sigma\sigma\sigma$
HEAVY		*!
LEFT-ALTERNATING	**	*

ALTERNATING in LB) into direct conflict for the first time. In producing forms from the test set, participants were required to determine the placement of stress and thus had to confront the conflict between HEAVY and LEFT-ALTERNATING (RIGHT-ALTERNATING in LB). All of the test words were of three types. Two of these types had five syllables and tested the critical transitive relation for the trained language only. The third type had seven syllables, with the syllable “ton” in the fourth position (e.g., so la ti ton do re mi), and tested transitivity in both LA and LB. Table 9 illustrates the conflict between HEAVY and LEFT-ALTERNATING that arises in the evaluation of this form in LA. The first candidate succeeds in stressing the heavy syllable, but at the expense of a LEFT-ALTERNATING violation that stems from its failure to stress the third syllable. This candidate incurs a second violation of LEFT-ALTERNATING due to its failure to stress the fifth syllable, but the training data already establish the dominance of penult stress in this situation (see Table 4). The second candidate fails to stress the heavy syllable, but succeeds in placing stress on the first and third syllable, as required by LEFT-ALTERNATING. The single violation of LEFT-ALTERNATING is due to the absence of stress on the fifth syllable, as with the other candidate. The ranking between HEAVY and LEFT-ALTERNATING will determine the choice between these two candidates. The independent rankings HEAVY  $\gg$  PENULT and PENULT  $\gg$  LEFT-ALTERNATING have been established in the training set, and transitivity gives HEAVY  $\gg$  LEFT-ALTERNATING. Thus, the stress pattern expected on the basis of transitive ranking is the first candidate in Table 9,  $\sigma\sigma\sigma\sigma_{\text{H}}\sigma\sigma\sigma$ . The stress pattern of the second candidate im-

TABLE 10

Five-Syllable Test Item for LA: Conflict between  
HEAVY and LEFT-ALTERNATING

$\sigma\sigma_{\text{H}}\sigma\sigma\sigma$	$\sigma\sigma\sigma_{\text{H}}\sigma\sigma\sigma$	$\sigma\sigma_{\text{H}}\sigma\sigma\sigma$
HEAVY		*!
LEFT-ALTERNATING	**	*

plies the dominance of LEFT-ALTERNATING over HEAVY, defying transitivity.

Since the heavy syllable is in the middle of the seven-syllable test words, the same item,  $\sigma\sigma\sigma\sigma_{\text{H}}\sigma\sigma\sigma$ , serves to test transitivity in LB, this time putting HEAVY into conflict with RIGHT-ALTERNATING.

The five-syllable test type with “ton” in the second position (e.g., mi ton fa so la) tested transitivity in LA only. Table 10 demonstrates the conflict between HEAVY and LEFT-ALTERNATING that arises in the evaluation of this form in LA.

The five-syllable test type with “ton” in the fourth position (e.g., so la ti ton do) tested transitivity in LB only, as shown in Table 11. The possible test words with their transitive stress patterns are shown in Table 12.

Notice that if transitivity applies, the five-syllable transitive test items require the participant to begin with an unstressed syllable in LA or end with an unstressed syllable in LB. In order to give transitivity a stringent test, there were no training stimuli that began with an unstressed syllable in LA or ended with an unstressed syllable in LB (see Table 8). This meant that the word type of three CV syllables ( $\sigma\sigma\sigma$ ) was not included in the training, because both LA and LB stress these words on the middle syllable. Other word types that begin

TABLE 11

Five-Syllable Test Item for LB: Conflict between  
HEAVY and RIGHT-ALTERNATING

$\sigma\sigma\sigma\sigma_{\text{H}}\sigma$	$\sigma\sigma\sigma\sigma_{\text{H}}\sigma$	$\sigma\sigma\sigma\sigma_{\text{H}}\sigma$
HEAVY		*!
RIGHT-ALTERNATING	**	*

TABLE 12

LA and LB Test Words with the Transitive Stress Patterns Indicated

LA five-syllable test words	LB five-syllable test words	Seven-syllable test words with LA stress pattern	Seven-syllable test words with LB stress pattern
do tón re mí fa	do ré mi tón fa	dó re mi tón fa só la	do ré mi tón fa so lá
re tón mi fá so	re mí fa tón so	ré mi fa tón so lá ti	re mí fa tón so la tí
mi tón fa só la	mi fá so tón la	mí fa so tón la tí do	mi fá so tón la ti dó
fa tón so lá ti	fa só la tón ti	fá so la tón ti dó re	fa só la tón ti do ré
so tón la tí do	so lá ti tón do	só la ti tón do ré mi	so lá ti tón do re mí
la tón ti dó re	la tí do tón re	lá ti do tón re mí fa	la tí do tón re mi fá
ti tón do ré mi	ti dó re tón mi	tí do re tón mi fá so	ti dó re tón mi fa só

with unstressed syllables in LA or end with unstressed syllables in LB involve HEAVY interacting with LEFT-ALTERNATING and hence they were already excluded from the training stimuli because they place the critical (transitive) constraints in conflict.

#### Procedure

The experiment had four parts: a repetition phase and a production phase during training, four internal tests spaced throughout the production phase, and a final test.

*Training.* Participants were told that they would be learning to pronounce words in a new language and that this language differed from English in two ways. First, the syllables that made up the words were those from the musical scale: do, re, mi, fa, so, la, and ti, as well as the syllable "ton." The second way that the new language differed from English, they were told, was where it accented or stressed the words.

Participants were given instructions for all four parts of the experiment before they began and were reminded of the appropriate instructions as they reached each section. The words were presented visually on a computer screen and the sessions were tape recorded. During the repetition phase of the experiment, the experimenter said each word with its correct pronunciation after it appeared on the computer screen. Stressed syllables were differentiated from unstressed syllables on this and all training items, most notably through the use of vowel reduction in unstressed syllables. Because the participants and the experimenter were native English speakers, other mark-

ers of stress typical of English such as a shorter vowel duration in unstressed syllables and a pitch peak on stressed vowels were present as well. After seeing and hearing one example in the instruction phase, participants repeated each of the next 125 words after the experimenter, while also viewing it on the computer screen. These consisted of all the words demonstrating HEAVY  $\gg$  PENULT (PEN-INITIAL), PENULT (PEN-INITIAL)  $\gg$  LEFT-ALTERNATING (RIGHT-ALTERNATING), and CLASH  $\gg$  HEAVY except those to be used in the internal tests. Each of the words was presented to the participants twice during the repetition phase at randomly determined positions.

The rest of the training consisted of a production task, in which participants saw a word on the computer screen and had to pronounce it. If they pronounced the word correctly, the experimenter told them that they said it correctly and presented the next word. If they pronounced the word incorrectly, they were told the correct pronunciation and had to correctly produce it twice before moving to the next word. The words in the production phase consisted of all of the words used in the repetition phase, as well as three words each from two of the types used in the internal tests but that were not themselves selected for use in one of the internal tests. As in the repetition phase, participants saw each word twice at randomly determined positions in the training. All of the words in both the repetition and production training phases were presented individually and in a random order for each subject. None of the words in either training phase directly demonstrated the domi-

nance relation between the constraints HEAVY and LEFT-ALTERNATING (RIGHT-ALTERNATING in LB).

*Internal tests.* Four internal tests were presented during the production phase of the training session to assess learning and to acclimate the participants to producing the words without feedback. Each test consisted of four words presented visually for production by the participant. None of the words in the internal tests had been seen by the participants during training. Two of the words were of familiar types, meaning that although the participants had not seen the exact words in the internal tests, they had seen words of that type during training, and two were of novel types (LA:  $\sigma_H\sigma\sigma\sigma\sigma$ ,  $\sigma\sigma\sigma\sigma\sigma_H$ ; LB:  $\sigma\sigma\sigma\sigma\sigma_H$ ,  $\sigma_H\sigma\sigma\sigma\sigma$ ). Each type tested either the HEAVY  $\gg$  PENULT (PEN-INITIAL in LB) domination order or the PENULT (PEN-INITIAL in LB)  $\gg$  LEFT-ALTERNATING (RIGHT-ALTERNATING in LB) domination order. The internal tests were produced by randomly selecting four words from each of the four internal test types. One word from each type was randomly presented in each internal test. Internal tests were placed after the 41st, 73rd, 105th, and 138th items in the production phase of the training session. Participants did not receive feedback on their productions during the internal tests.

*Final test.* The participants in both languages were asked to pronounce the same word types during the 15-item final test. Five of the items tested for transitivity in both languages (seven-syllable test words; e.g., fa so la ton ti do re), 5 tested for transitivity in LA only (five-syllable test words; e.g., re ton mi fa so), and 5 tested for transitivity in LB only (five-syllable test words; e.g., do re mi ton fa). Five of seven possible test words of each type were randomly selected for each participant, such that each saw 10 test items (5 five-syllable and 5 seven-syllable) that put the two critical constraints into conflict for the language they were trained on and 5 test items that put the two critical constraints into conflict for the language they were not trained on. The final test came directly after the fourth internal test, and participants were asked, simi-

larly to the internal tests, to produce the words without any feedback.

### Results

The recordings were analyzed to determine which syllables were stressed in the test words. The main coding was done by the first author. Since all of the participants and coders were native English speakers, coders listened for any and all of the markers of stress in English. These include vowel reduction in unstressed syllables, vowel shortening in unstressed syllables, and pitch peaks on stressed vowels. Coding reliability was fairly good. We did two reliability checks. The first used a subset of the data from the final test phase of Experiment 3, a study that had the same test materials as Experiment 1. A second independent coder<sup>3</sup> agreed with the main coder's assignment of stress values to syllables 93% of the time (based on 60 words containing a total of 340 syllables). The second check involved data from a study that is not presented here, but used similar procedures and the same materials as Experiment 1. (Its findings were similar to those of Experiment 1). Another independent second coder agreed with the main coder's assignments 96% of the time (based on 336 words containing 1904 syllables). Incorrect responses during training were tracked by the experimenter during the experimental session. The percentage of correct responses during the production phase of training was 54% overall and 64% during the last quarter of the production phase of training.

Responses for the critical test items were each put into one of seven categories for both five- and seven-syllable test words. The first

<sup>3</sup> All coding, including the reliability checks, was done on responses that were blocked by speakers. This makes it easier to hear the intended stress patterns, because speakers differ in speech rate and other factors that impact timing or vowel reduction. A consequence of this blocking is that the second coders, even though they were not explicitly informed of the condition in which each speaker participated, can easily hear differences that indicate the training language. Thus, the reliability check is not blind with respect to training language. However, the check on the coding from Experiment 3 was blind with respect to the key condition of that experiment, whether seven-syllable items were included or not in the training set.

TABLE 13

Mean Percentages of Response Types Produced for Five- and Seven-Syllable Test Items (*SD*)

	Transitive	Template A (critical nontransitive)	Template B	Template C	Opposite language transitive	CLASH violations	Other
<b>LA</b>							
Five-syllable	36 (28)	32 (36)	2 (6)	2 (6)	NA <sup>a</sup>	16 (30)	12 (19)
Seven-syllable	24 (26)	34 (25)	12 (17)	8 (10)	2 (6)	10 (17)	10 (14)
<b>LB</b>							
Five-syllable	24 (28)	16 (16)	10 (17)	38 (29)	NA	8 (10)	4 (8)
Seven-syllable	32 (19)	24 (32)	6 (10)	10 (14)	0 (0)	18 (20)	10 (11)

<sup>a</sup> Opposite language Transitive responses are not possible on the five-syllable critical test items for a given language since the five-syllable items only test transitivity for one language or the other. For example, the five-syllable critical test items for LA had the form  $\sigma\sigma_H\sigma\sigma\sigma$ . This cannot be given a transitive pattern for LB regardless of how it is stressed because the test item's "ton" is in the wrong place. Opposite language responses were seen on five syllable items only for those items that tested transitivity in the opposite language. Since responses to those items were not included in Table 13, opposite language transitive responses for five-syllable items are not shown even though they occurred in both languages.

category, *transitive responses*, were those responses that perfectly followed the grammar of the language the participants had been trained in. Collapsing over languages, this category contained 29% of the responses. The second category, *opposite language transitive responses*, consisted of responses that would have been transitive responses had the participants been trained in the other language. This category contained 6% of the responses on the test items allowing such responses. The third category, *CLASH violations*, containing 13% of the responses, consisted of all of the responses that stressed two or more contiguous syllables. The next three categories consist of what will be referred to as "template" responses. These responses had stress patterns identical to those of words of the same syllable length seen during training. There were three possible stress patterns seen during training for five- and for seven-syllable words from which templates could be derived. For both the five- and seven-syllable test words, there were three possible templates: A, B, and C. In *template-A* responses, there are two adjacent unstressed syllables before the penultimate syllable (in LA) or after the pen-initial syllable (in LB) (**LA:**  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ,  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ; **LB:**  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ,  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ). These patterns accounted for 26.5% of the responses. *Template-B* re-

sponses had strictly alternating stress patterns beginning on the first syllable and ending on the final syllable in both LA and LB (**LA:**  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ,  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ; **LB:**  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ,  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ). *Template-B* patterns occurred on 7.5% of the responses. They may have been relatively infrequent because five-syllable items with the *Template B* stress pattern were used only in the internal tests and therefore could be used only to model stress patterns during the later tests to the degree that they were produced correctly during the internal tests. Finally, *template-C* responses were characterized by alternating stress with two adjacent unstressed syllables at either the end of the word (LA) or the beginning of the word (LB) (**LA:**  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ,  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ; **LB:**  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ,  $\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}\underline{\quad}$ ). *Template-C* accounts for 14.5% of the responses. All of the template responses are suboptimal to the transitive response since each violates the highly ranked constraint HEAVY in favor of satisfying lower ranked alignment constraints. The final category, *Other responses* (9%), consisted of any other responses to the critical test items as well as responses for which the main coder was unsure of the stress pattern produced for that token. The percentages of each type of response are given in Table 13 broken down by language and test type.

All statistical tests evaluating the transitivity hypotheses in this and subsequent experiments consisted of planned comparisons involving the number of transitive responses given and how often some other kind of response occurred. Because these numbers are not independent, we used a nonparametric test appropriate for paired data, the Wilcoxon. To gain power, we computed the following tests including participants trained in both languages. We do, however, present the means broken down by language group.

Two tests of transitivity were carried out, a weak test and a strong test. The weak test was a comparison between the number of transitive responses and opposite language transitive responses. More test items were produced with the transitive stress pattern (29%) than with the opposite language transitive stress pattern (6%). This difference was significant ( $Z = -3.64, p < .05$ ). The weak test shows that participants did learn something from their training and whatever they learned increased their tendency to give the correct stress pattern on the test of transitivity. Any such tendency cannot be attributed solely to their knowledge of English or any other previous knowledge.

The strong test for transitivity was a comparison between the number of transitive responses produced and the number of template-A responses produced. This comparison is crucial because template-A responses (LA:  $\dot{\_} \_ \dot{\_} \_ \dot{\_}$ ,  $\dot{\_} \_ \_ \dot{\_} \_ \dot{\_}$ ; LB:  $\_ \dot{\_} \_ \dot{\_} \_ \dot{\_}$ ,  $\_ \_ \dot{\_} \_ \dot{\_} \_ \dot{\_}$ ) are what would be expected on the test items if the explicitly demonstrated constraint relations are learned, yet LEFT-ALTERNATING(RIGHT-ALTERNATING) was satisfied over HEAVY, thus violating transitivity. Template-A responses will, therefore, also be referred to as the *critical nontransitive responses*. This test did not reach significance at the .05 level for either the five-syllable test words (transitive = 30%, critical nontransitive = 24%),  $Z = -0.60, p > .05$ , or for the seven-syllable test words (transitive = 28%, critical nontransitive = 29%),  $Z = -0.04, p > .05$ . Finally, no differences were found in the number of transitive responses produced

by the two language groups,  $t(18) = .21, p > .05$ .

### Discussion

The results of Experiment 1 give support for a transitive constraint system under the weak test of transitivity. Participants were nearly five times as likely to give the transitive response for the language they were trained in than the one they were not trained in. This, combined with the fact that there were no differences between the number of transitive responses produced by each language group, demonstrates that the responses were due to the training the participants received rather than their previous knowledge of English or some idiosyncrasy of one of the languages.

The strong test of transitivity, however, does not support a strictly ranked constraint system. In order for a transitive relationship to be demonstrated, participants must show that, at least on a given trial, the trained relationships HEAVY  $\gg$  PENULT(PEN-INITIAL) and PENULT(PEN-INITIAL)  $\gg$  LEFT-ALTERNATING(RIGHT-ALTERNATING) were followed. The only responses that correctly demonstrate both of these relations are the transitive responses themselves and the critical nontransitive responses (template-A). When these relationships were satisfied, however, the transitive relationship between the constraints HEAVY and LEFT-ALTERNATING(RIGHT-ALTERNATING) was satisfied only 56% of the time for five-syllable words (30% transitive versus 24% critical nontransitive) and 49% of the time for seven-syllable words (28% transitive versus 29% critical nontransitive). If participants failed to apply transitivity, they would be equally likely to produce a form showing HEAVY  $\gg$  LEFT-ALTERNATING(RIGHT-ALTERNATING) and one showing LEFT-ALTERNATING(RIGHT-ALTERNATING)  $\gg$  HEAVY. Hence, by chance, one would expect these percentages to be 50%, and the obtained percentages do not differ significantly from this. The results suggest that, when faced with a constraint conflict that they had not experienced, participants chose to satisfy each of the constraints an equal number of times.

The failure of the strong test of transitivity could indicate that speakers do not use a ranked constraint system to determine the pronunciation of novel forms. Alternatively, 30 min of training may not have been enough to develop a solid constraint system. A small experiment was therefore carried out to assess the effects of extended training in the two languages.

#### *Performance under Extended Training*

To determine if insufficient learning of the artificial languages was responsible for the results of Experiment 1, an experiment in which participants were trained in the languages for 4 days before final testing was carried out. There were internal tests analogous to those of Experiment 1 after the first, second, and third days. The experiment had a procedure very similar to that of Experiment 1. The same training words and constraint systems used in Experiment 1 were used here. The only difference in the stimuli was the addition of a six-syllable test type for each language (**LA:**  $\sigma\sigma_H\sigma\sigma\sigma\sigma$ ; **LB:**  $\sigma\sigma\sigma\sigma\sigma_H\sigma$ ). Because of the addition of these test types, the final test contained 25 items.

Four participants were each trained for 4 days. All four participants were 100% correct on the internal test that occurred after the third day of training and on the training session of day 4, thus achieving the goal of excellent learning of the trained stimulus types. The test words were presented for the first time at the end of the fourth session. As in Experiment 1, there were more transitive (22%) than opposite language transitive responses (0%). Even though there were no opposite language transitive responses, this comparison did not quite reach statistical significance due to the small number of participants,  $Z = -1.84$ ,  $p = .07$ .

The strong tests of transitivity, comparing the number of transitive responses with the number of critical nontransitive (template-A) responses, were, as before, unsuccessful for both five-syllable test items (transitive = 5%, critical nontransitive = 40%),  $Z = 1.63$ ,  $p > .05$ , and seven-syllable test items (transitive = 45%, critical nontransitive = 35%),  $Z = .00$ ,  $p > .05$ . The strong test also showed no preference for transitive responses for the new six-syllable test

items (transitive = 15%, critical nontransitive = 35%),  $Z = 1.11$ ,  $p > .05$ .

Increasing the amount of training thus did not increase the likelihood of transitive responses. In fact, only 22% of the responses in the extended training experiment had the transitive stress pattern compared to 29% of the responses in Experiment 1. As in Experiment 1, there was no difference between the number of transitive and critical nontransitive (template-A) responses. On the responses that satisfied all the constraint relationships demonstrated during training, transitivity was followed only 37% of the time. As mentioned earlier, this could be due to the absence of a strictly ranked constraint system.

However, examination of all of the nontransitive responses suggests another interpretation. In the extended training experiment, 62% of the responses on items that test for transitivity had identical stress patterns to words of the same syllable length that were seen during training. That is, they fit one of the template patterns. In Experiment 1, such patterns comprised 49% of the responses. Thus, the participants may have developed metrical templates for words with a certain number of syllables (regardless of the CV structure of the syllable) and then applied these templates to the novel forms.

Evidence that stress patterns can be stored separately from a word's segmental information can be seen in studies of the tip-of-the-tongue phenomenon (Brown & McNeill, 1966; Meyer & Bock, 1992; Rubin, 1975) and implicit priming studies (Roelofs & Meyer, 1998). The tip-of-the-tongue state is one in which a speaker has an intense feeling of knowing a word, without being able to produce it. Quite often, even though the word's segments can not be retrieved, the speaker is able to correctly report the number of syllables the word contains and give evidence of knowing word's stress pattern as well. The basis of this knowledge could be metrical templates.

Further evidence that a word's metrical structure may be stored in a form independent not only of segmental information but also of CV structure is given by Roelofs and Meyer (1998). In a series of implicit priming experiments,

Roelofs and Meyer presented participants with Dutch word pairs in sets in which the target words were either held constant or varied with regard to segmental overlap, the number of syllables, main stress position, and CV structure. The participants were then required to quickly recall the second word of each of the word pairs from the set they saw. Roelofs and Meyer showed that priming due to shared word onsets occurred only when the number of syllables and their stress patterns were held constant. They interpreted this as evidence of independently stored metrical frames, or what is referred to here as metrical templates. Interestingly, CV structure did not play a role in priming, suggesting that CV structure is not encoded in the metrical frames. This could explain why a metrical template from a word with a different CV structure but same number of syllables could be associated to the test words in Experiment 1. Recently, Levelt, Roelofs, and Meyer (1999) amended the proposal of Roelofs and Meyer (1998). They kept the idea that there are stored metrical frames that lack CV structure, but only for lexical items that do not follow the default stress pattern for the language.

These studies support the idea that known words are associated with metrical templates (lacking CV structure) and that words using the same pattern activate a common template. The existence of shared templates, across words, suggests that templatic information can be transferred to novel items that match on syllable length. Thus, a speaker knows how to stress the new string fá so lá ti because the template ' \_ \_ ' \_ is stored.

These stored templates may have been responsible for the failure of the strong tests of transitivity in Experiment 1. Since none of the training words in these experiments put HEAVY and LEFT-ALTERNATING (RIGHT-ALTERNATING) into conflict, any five- and seven-syllable training words that did not put HEAVY and PENULT (PEN-INITIAL) into conflict necessarily had the same stress pattern as that of the critical nontransitive responses. The critical nontransitive responses thus have the unfair advantage of being associated with frequently used templates, whereas the transi-

tive responses, by design, were completely novel patterns and thus lacked templates arising from training. Therefore, the transitive and critical nontransitive responses may be produced by two different processes. The former is produced by a system of violable constraints and the later by a system associating words of a certain syllable length to stored metrical templates. Experiment 2 was carried out to examine this hypothesis.

## EXPERIMENT 2

If output forms are determined by two different processes, a ranked constraint system and a system of metrical templates, then if one of these systems can be rendered ineffectual, output forms will have to be determined by the other system. Experiment 2 attempts to do this by adding a condition to the procedures used in Experiment 1 in which participants are never presented with seven-syllable words during training. Because they never see or hear any seven-syllable words during training, the participants in this *holdout* condition will not have any words of this length in the artificial languages from which to build templates. They should, therefore, show more evidence of transitivity on the strong test with the seven-syllable test items than with the five-syllable test items or than the seven-syllable and five-syllable test items for the participants in the *full*, or non-holdout, condition.

### *Method*

#### *Participants*

Forty native English-speaking undergraduate students from the University of Illinois participated in exchange for either credit toward a course requirement or \$5. All participants were randomly and equally assigned to one of two language conditions (LA and LB) and one of two training conditions (holdout condition and full condition).

#### *Stimulus Materials*

The same constraint systems and words used during the training of Experiment 1 were used here. The main difference between Experiments

TABLE 14  
Experiment 2, LA and LB Training Types

Language A			Language B		
CLASH >> HEAVY	HEAVY >> PENULT	PENULT >> LEFT- ALTERNATING	CLASH >> HEAVY	HEAVY >> PEN-INITIAL	PEN-INITIAL >> RIGHT- ALTERNATING
$\acute{\sigma}_H\sigma_H\acute{\sigma}\sigma$ (22/38)	$\acute{\sigma}\acute{\sigma}_H$ (21/38)	$\acute{\sigma}\sigma\acute{\sigma}\sigma$ (22/38)	$\acute{\sigma}\sigma\sigma_H\acute{\sigma}_H$ (22/38)	$\acute{\sigma}_H\acute{\sigma}\acute{\sigma}$ (21/38)	$\acute{\sigma}\sigma\sigma\sigma$ (22/38)
$\acute{\sigma}\sigma_H\acute{\sigma}_H\acute{\sigma}\acute{\sigma}$ (21/36)	$\acute{\sigma}\sigma\acute{\sigma}\acute{\sigma}_H$ (22/38)	$\acute{\sigma}_H\sigma\sigma\sigma$ (22/38)	$\acute{\sigma}\sigma\acute{\sigma}_H\sigma_H\acute{\sigma}$ (21/36)	$\acute{\sigma}_H\sigma\acute{\sigma}\acute{\sigma}\acute{\sigma}$ (22/38)	$\acute{\sigma}\sigma\sigma\acute{\sigma}_H$ (22/38)
	$\acute{\sigma}\sigma\acute{\sigma}_H\sigma\sigma$ (23/38)	$\acute{\sigma}\sigma\acute{\sigma}\sigma\sigma\sigma$ (28/0)		$\sigma\sigma\acute{\sigma}_H\acute{\sigma}\acute{\sigma}$ (23/38)	$\acute{\sigma}\sigma\sigma\acute{\sigma}\acute{\sigma}\acute{\sigma}$ (28/0)
	$\acute{\sigma}\sigma\acute{\sigma}\acute{\sigma}_H\sigma\sigma$ (27/0)	$\acute{\sigma}\sigma\sigma_H\sigma\sigma\sigma$ (28/0)		$\sigma\sigma\acute{\sigma}_H\sigma\acute{\sigma}\acute{\sigma}\acute{\sigma}$ (27/0)	$\acute{\sigma}\sigma\sigma\acute{\sigma}_H\acute{\sigma}\acute{\sigma}$ (28/0)
	$\acute{\sigma}\sigma\acute{\sigma}\sigma\sigma\acute{\sigma}_H$ (28/0)			$\acute{\sigma}_H\sigma\sigma\acute{\sigma}\acute{\sigma}\acute{\sigma}$ (28/0)	

Note. Number of tokens in full condition/number of tokens in holdout condition.

1 and 2 was the addition of the holdout condition within each language. Since the holdout condition eliminated a number of the training items, two internal test types held out from training in Experiment 1 were used during training in Experiment 2 so that there would be a great enough variety in the training stimuli to demonstrate the critical constraint relationships. One of these demonstrated the HEAVY >> PENULT relationship in LA,  $\acute{\sigma}\sigma\acute{\sigma}\sigma_H$ , and the HEAVY >> PEN-INITIAL relationship in LB,  $\acute{\sigma}_H\sigma\acute{\sigma}\acute{\sigma}\acute{\sigma}$ . The other demonstrated the PENULT >> LEFT-ALTERNATING relationship in LA,  $\acute{\sigma}_H\sigma\sigma\sigma\sigma$ , and the PEN-INITIAL >> RIGHT-ALTERNATING relationship in LB,  $\acute{\sigma}\sigma\sigma\sigma\acute{\sigma}_H$ . In addition, in order to keep the same relative number of training types that demonstrate the two critical constraint relationships and to maintain the same number of five- and seven-syllable training types, the LA types  $\acute{\sigma}_H\sigma\sigma$  and  $\acute{\sigma}_H\sigma\acute{\sigma}\sigma\sigma$  and the LB types  $\sigma\sigma\acute{\sigma}_H$  and  $\sigma\sigma\acute{\sigma}\sigma\acute{\sigma}_H$  were not used. The training types and the number of tokens used in Experiment 2 are given in Table 14.

To keep Experiments 1 and 2 comparable, and to accustom the participants to reading these words without any feedback, internal tests were performed in a manner similar to that of Experiment 1. The internal test items, in contrast to Experiment 1, were training items that participants had seen previously. The same final test items used in Experiment 1 were used during the final test of Experiment 2.

Procedure

Experiment 2 was run in an identical fashion to Experiment 1 with one exception. In contrast to Experiment 1, 10 final test lists were randomly generated, and each was given to one participant in each of the four conditions (LA full, LA holdout, LB full, LB holdout). Thus, one participant from each of the groups was tested with the same 15 test items in the same random pattern.

Results

The percentage of correct responses during the production phase of training was 69% overall and 72% during the last quarter of the production phase of training. For the test items, there were no differences in the overall number of transitive responses produced by the two language groups,  $t(38) = 0.65, p > .05$ . Again, there were more transitive responses (full condition = 22%; holdout condition = 31%) than opposite language transitive responses (full condition = 3%; holdout condition = 3%), full:  $Z = -3.16, p < .05$ ; holdout:  $Z = -3.46, p < .05$ , demonstrating that the transitive responses given by both language groups can be attributed to some aspect of the training. Percentages of response types are given in Table 15.

The major comparison of interest in Experiment 2 was the strong test of transitivity, comparing the transitive responses with the critical nontransitive (template-A) responses. If the hy-

TABLE 15

Percentages of Response Types Produced for Five- and Seven-Syllable Test Items (*SD*)

	Transitive	Template A (critical nontransitive)	Template B	Template C	Opposite language transitive	CLASH violations	Other
LA							
Full							
Five-syllable	32 (37)	24 (31)	20 (30)	12 (32)	NA	10 (17)	2 (6)
Seven-syllable	6 (14)	22 (27)	26 (31)	0 (0)	2 (6)	30 (41)	14 (27)
Holdout							
Five-syllable	36 (37)	18 (14)	20 (28)	2 (6)	NA	16 (18)	8 (10)
Seven-syllable	20 (42)	14 (21)	18 (29)	2 (6)	0 (0)	26 (25)	20 (25)
LB							
Full							
Five-syllable	14 (25)	26 (21)	4 (8)	38 (27)	NA	2 (6)	16 (25)
Seven-syllable	34 (33)	6 (10)	6 (10)	16 (18)	0 (0)	4 (13)	34 (23)
Holdout							
Five-syllable	16 (28)	32 (25)	8 (10)	22 (27)	NA	10 (14)	12 (21)
Seven-syllable	52 (25)	2 (6)	0 (0)	16 (16)	0 (0)	8 (10)	22 (24)

pothesis of competing constraint and template systems is correct, we would expect to see more evidence of transitivity on the strong test when there were no training words from which to develop templates, as was the case with the seven-syllable test items in the holdout condition. However, no more success than in Experiment 1 should be expected for the strong tests of transitivity on the five-syllable test items for the holdout condition or for the five- and seven-syllable test items for the full condition since words were available during training in all of these cases from which to build templates. As expected, more transitive responses (36%) were produced for the seven-syllable test items than critical nontransitive responses (8%) in the holdout condition  $Z = -2.37, p < .05$ . There was, however, no significant difference between the number of transitive responses (20%) and critical nontransitive responses for the seven-syllable items in the full condition (14%)  $Z = 0.71, p > .05$ . The other comparisons of interest were between the five-syllable transitive responses and the five-syllable critical nontransitive responses produced in both the holdout and full conditions. Similarly to Experiment 1, where five-syllable training items with the crit-

ical nontransitive stress patterns were given, neither the holdout or full conditions produced significantly more transitive (holdout = 26%; full = 23%) than critical nontransitive responses (holdout = 25%; full = 25%), holdout:  $Z = 0.05, p > .05$ ; full:  $Z = .22, p > .05$ . The evidence for transitivity in the holdout condition for seven-syllable items came from both languages. In LA, transitive responses increased from 6 to 20% from the full to the holdout conditions, while critical nontransitive responses decreased from 22 to 14%. In LB, holding out seven-syllable training items increased transitive responses from 34 to 52% and decreased critical nontransitive responses from 6 to 2%.

### Discussion

When the number of transitive responses is compared to the number of critical nontransitive responses for the five-syllable test items, there are no differences for either the full or holdout conditions. This might be expected based on Experiment 1 since both groups saw five-syllable training words and therefore were able to form templates for these items. When the seven-syllable test items are considered, there was

again no difference for the full condition, with transitivity applying on 59% of the responses for which all of the demonstrated constraint relationships were followed (20% transitive versus 14% critical nontransitive). This is most likely due to the fact that the participants in this group saw seven-syllable words during training. However, when the seven-syllable items from the holdout condition are considered, there are many more transitive responses produced than critical nontransitive responses. When all of the constraint relationships directly demonstrated during training were followed, transitivity was applied 82% of the time (36% transitive versus 8% critical nontransitive). This difference conforms to the expected results since there were no seven-syllable words seen during training from which to develop templates. The results of Experiment 2 thus support the hypothesis of a two-process system, one using ranked constraints and the other associating inputs to stored templates.

### EXPERIMENT 3

Experiment 2 provided evidence of a transitive constraint system as well as a system that matches input to previously stored templates. A third experiment using the same basic design was run to replicate the findings of Experiment 2. Experiment 3 differed from the previous one by eliminating the production phase of training so participants learned the artificial language only by hearing and repeating the training items. This greatly simplified training from the participant's perspective. Repeating what is heard can be done largely without error. This nearly error-free training procedure was chosen in an effort to reduce the likelihood of participants employing conscious strategies to learn the experimental languages.

#### *Method*

##### *Participants*

Twenty native English-speaking undergraduates from the University of Illinois participated in exchange for either credit toward a course requirement or \$5. Five participants were randomly assigned to each of the four experimental

conditions (LA full, LA holdout, LB full, LB holdout).

##### *Stimulus Materials*

The same training types used in Experiment 2 were used, as well as the LA type  $\acute{\sigma}_H\sigma\sigma$  and the LB type  $\sigma\sigma\acute{\sigma}_H$ . These types demonstrate the HEAVY  $\gg$  PENULT(PEN-INITIAL) relationship and were added to partly offset an increased frequency of tokens demonstrating the PENULT  $\gg$  LEFT-ALTERNATING (RIGHT-ALTERNATING) constraint relationship. The number of words given during training that demonstrate this relationship were increased after a pilot study using the procedure employed in Experiment 3 indicated that participants had slightly more trouble learning the PENULT  $\gg$  LEFT-ALTERNATING (RIGHT-ALTERNATING) constraint relationship than the others. To remedy this, we doubled the number of times each of the words demonstrating this relationship was shown. This provided only slightly more demonstrations of the PENULT  $\gg$  LEFT-ALTERNATING (RIGHT-ALTERNATING) relationship than the HEAVY  $\gg$  PENULT relationship since there were only five types demonstrating PENULT  $\gg$  LEFT-ALTERNATING (RIGHT-ALTERNATING) and six that demonstrated HEAVY  $\gg$  PENULT. Table 16 shows the training types and their frequencies.

##### *Procedure*

All materials, including auditory presentation of the instructions, training words, and feedback, were presented by a computer. When the participant indicated that she or he was ready to begin, the experimenter pressed a key to begin the instructions. The direction "please listen carefully to the following instructions" appeared on the screen, and an audio file containing the prerecorded instructions began. The words on the screen were replaced by the example word "fa so la ti do" at a point in the instructions demonstrating what a word in the language would look and sound like. The only difference in the instructions for LA and LB was the sound file used for demonstration of the pronunciation for the example word. There were no differences between the instructions for

TABLE 16  
Experiment 3, LA and LB Training Types

Language A			Language B		
CLASH >> HEAVY	HEAVY >> PENULT	PENULT >> LEFT- ALTERNATING	CLASH >> HEAVY	HEAVY >> PEN-INITIAL	PEN-INITIAL >> RIGHT- ALTERNATING
$\acute{o}_H\sigma_H\acute{s}\acute{s}\acute{o}$ (14/14)	$\acute{o}_H\sigma\sigma$ (14/14)	$\acute{o}\sigma\acute{s}\acute{o}$ (28/56)	$\sigma\acute{s}\sigma_H\acute{o}_H$ (14/14)	$\sigma\acute{s}\acute{o}_H$ (14/14)	$\sigma\acute{s}\sigma\sigma$ (28/56)
$\acute{o}\sigma_H\acute{o}_H\sigma\acute{s}\acute{o}$ (14/14)	$\acute{o}\sigma\acute{o}_H$ (14/14)	$\acute{o}_H\sigma\sigma\acute{s}\acute{o}$ (28/56)	$\sigma\acute{s}\sigma\acute{o}_H\sigma_H\acute{o}$ (14/14)	$\acute{o}_H\sigma\acute{o}$ (14/14)	$\sigma\acute{s}\sigma\sigma_H$ (28/56)
	$\acute{o}\sigma\sigma\acute{o}\acute{o}_H$ (14/28)	$\acute{o}\sigma\sigma\acute{o}\sigma\acute{s}\acute{o}$ (28/0)		$\acute{o}_H\sigma\acute{s}\acute{o}\acute{s}$ (14/28)	$\sigma\acute{s}\sigma\sigma\sigma\acute{s}$ (28/0)
	$\acute{o}\sigma\acute{o}_H\sigma\sigma$ (14/28)	$\acute{o}_H\sigma\acute{s}\sigma\sigma\acute{s}\acute{o}$ (28/0)		$\sigma\sigma\acute{o}_H\sigma\acute{s}$ (14/28)	$\sigma\acute{s}\sigma\sigma\sigma\sigma\acute{s}$ (28/0)
	$\acute{o}\sigma\acute{o}\sigma\acute{o}_H\sigma\sigma$ (14/0)			$\sigma\sigma\acute{o}_H\sigma\acute{s}\acute{o}\acute{s}$ (14/0)	
	$\acute{o}\sigma\acute{s}\sigma\acute{s}\sigma\acute{o}_H$ (14/0)			$\acute{o}_H\sigma\acute{s}\sigma\acute{s}\sigma\acute{s}$ (14/0)	

*Note.* Number of tokens in full condition/number of tokens in holdout condition.

the holdout and non-holdout conditions. At the end of the instructions, the participant was given the opportunity to ask questions. The experimenter then pressed a key to begin the training phase of the experiment.

*Training.* To add an extra measure of control, one token of each word for both languages was digitally recorded and stored on the computer. Although the participants in the holdout condition did not encounter all of the possible training words, the words that both the holdout and non-holdout conditions did hear had exactly the same pronunciation.

At the beginning of the training session, the participants saw on the computer screen and heard the computer “say” “please repeat the following words.” The participants then had to repeat the next 224 words. The words were presented visually for 1000 ms, and then the words were presented auditorily by the computer while the word remained on the screen. If the participant repeated the word correctly, the experimenter pressed a button and the computer said “correct.” The screen then went blank for 1000 ms and the next word was presented. If a word was pronounced incorrectly, the experimenter pressed a key that played an audio file saying “the correct pronunciation is” and then the word’s sound file was played again. The participants then had to repeat the word again. All training items were presented in random order.

### Test

After the training phase, participants heard and saw the words “please read the following words.” They then were presented with visual but not auditory representations of all 21 possible test words used in Experiments 1 and 2 (Table 6). After the participant produced each word, the experimenter pressed a key that cleared the screen for 1000 ms and presented the next test item. All test items were presented in random order.

### Results

Training words were produced with 99% accuracy. For the test items, no differences were found in the overall number of transitive responses between the two language groups,  $t(18) = -.71, p > .05$ . As in the other experiments, there were more transitive responses (full condition = 16%; holdout condition = 37%) than opposite language transitive responses (full condition = 1%; holdout condition = 1%), full:  $Z = -3.11, p < .05$ ; holdout:  $Z = -3.55, p < .05$ , demonstrating that the transitive responses given by both language groups can be attributed to some aspect of the training. Percentages of response types are given in Table 17.

As in Experiment 2, the major comparison of interest was the strong test of transitivity, comparing the transitive responses with the critical nontransitive (template-A) responses. Again,

TABLE 17

Percentages of Response Types Produced for Five- and Seven-Syllable Test Items (*SD*)

	Transitive	Template A (critical nontransitive)	Template B	Template C	Opposite language transitive	CLASH violations	Other
LA							
Full							
Five-syllable	23 (30)	26 (26)	11 (16)	9 (19)	NA	17 (31)	14 (20)
Seven-syllable	9 (13)	20 (31)	9 (13)	17 (26)	0 (0)	14 (18)	31 (42)
Holdout							
Five-syllable	23 (22)	14 (18)	26 (26)	17 (26)	NA	11 (12)	39 (18)
Seven-syllable	40 (23)	0 (0)	9 (13)	6 (8)	0 (0)	26 (31)	20 (13)
LB							
Full							
Five-syllable	20 (13)	14 (20)	6 (8)	23 (16)	NA	29 (20)	9 (8)
Seven-syllable	14 (18)	9 (19)	3 (6)	17 (16)	0 (0)	23 (16)	34 (13)
Holdout							
Five-syllable	17 (16)	37 (22)	0 (0)	14 (25)	NA	20 (22)	11 (19)
Seven-syllable	66 (28)	0 (0)	0 (0)	11 (19)	0 (0)	11 (19)	11 (26)

more transitive responses (53%) were produced for the seven-syllable test items than critical nontransitive responses (0%) in the holdout condition,  $Z = -2.82$ ,  $p < .05$ . Also as in Experiment 2, there was no significant difference between the number of transitive responses (11%) and critical nontransitive responses for the seven-syllable items in the full condition (14%),  $Z = 0.0$ ,  $p > .05$ . The other comparisons of interest were between the five-syllable transitive responses and the five-syllable critical nontransitive responses produced in both the holdout and full conditions. As before, neither the holdout or full conditions produced significantly more transitive (holdout = 20%; full = 21%) than critical nontransitive responses (holdout = 26%; full = 20%), holdout:  $Z = 0.539$ ,  $p > .05$ ; full:  $Z = -0.18$ ,  $p > .05$ . To a greater degree than in Experiment 2, evidence for transitivity in the holdout condition for seven-syllable items came from both languages. In LA, transitive responses increased from 9 to 40% from the full to the holdout condition, while critical nontransitive responses decreased from 20 to 0%. In LB, holding out seven-syllable training items increased transitive responses

from 14 to 66% and decreased critical nontransitive responses from 9 to 0%.

### Discussion

Experiment 3 provides strong evidence in support of the dual-system hypothesis of competing template and constraint systems. When participants were given words during training from which they could derive templates, as with the five-syllable words in both the full and holdout conditions and the seven-syllable words in the full condition, participants were just as likely to give a transitive response pattern as a template response pattern when they satisfied all of the constraint relationships directly demonstrated during training. When there were no words from which to build templates, however, as with the seven-syllable words in the holdout condition, the transitive stress patterns were given 100% of the time that the constraint relationships demonstrated during training, and needed for the transitive system, were correctly applied in the test words (53% transitive versus 0% critical nontransitive). The evidence of transitivity was stronger in this experiment than in the previous one. We believe that the changes in

the procedures were responsible. Participants learned the sound patterns in the artificial language simply by listening and repeating.

Although the results of this experiment were strongly in line with the expectations of the dual-system hypothesis, one may question whether our participants were learning something that corresponds to a linguistic stress system, as opposed to some sort of nonlinguistic rhythmic pattern. One way to test this is to examine the data of those participants who are most likely to be treating the task linguistically. If these participants do not show the predicted pattern of results, then we would have reason for concern. The participants who most clearly treated the task as that of learning a stress system were those who reduced stressless syllables, importing this property from their native language, English. The data of the seven participants who did reduce most clearly (2 LA full, 3 LB full, 0 LA holdout, 2 LB holdout) provided strong support for the dual-system hypothesis. There was no difference between the number of seven-syllable transitive and critical nontransitive responses produced by the participants who reduced in the full condition. Both types were produced at a rate of 14%. Participants who reduced in the holdout condition, however, produced 79% of the seven-syllable critical test items with the transitive stress pattern and none with the critical nontransitive pattern. Although these percentages do not reflect a large number of trials, they show that the speakers who marked stress most naturally had data consistent with that shown by the full set of participants.

### GENERAL DISCUSSION

Our three experiments tested the validity of a key assumption of Optimality Theory, namely, that the relation of dominance between constraints is subject to transitivity. Experiment 1 showed that responses reflecting transitivity did occur as a result of learning an artificial language. The strong test of transitivity, however, which directly compared transitive responses to the alternative critical nontransitive responses, did not yield evidence for the application of transitivity. Responses showing that Constraint A  $\gg$  Constraint C were just as likely as re-

sponses showing that C  $\gg$  A. Most of the erroneous responses given to test items, however, exhibited stress patterns identical to those of training words with the same number of syllables. This led to our dual-system hypothesis: Words are produced by two competing processes, one governed by a transitive constraint system and the other consisting of a set of stored metrical templates.

Experiment 2 tested the dual-system hypothesis by creating a condition in which seven-syllable test items could not be stressed from a template derived during training. As expected, evidence of transitivity was then found with the strong test. Specifically, 82% of the seven-syllable test responses in the holdout condition that exhibited the demonstrated constraint relations were consistent with transitivity. In the third experiment, which used a simplified training procedure, 100% of these responses were transitive. When items could be assigned stress from a template derived from training in all of the experiments, there was no unambiguous evidence for transitivity.

The idea that language forms are computed from a grammatical system as well as from some kind of memory system or lexicon is a common theoretical claim in psycholinguistics. Examples include dual-system approaches to regular and irregular morphology (e.g., Pinker, 1991), word reading (e.g., Coltheart, Curtis, Atkins, & Haller, 1993), and, most relevant to our research, Levelt et al.'s (1999) account of the production of regular and irregular metrical patterns.

The claim that output can come from a general rule-like process acting in concert or in competition with a memory-based system is not restricted to the language domain. Logan's (1988) theory of automaticity proposes that responses in many tasks arise from either an algorithm (analogous to a grammar) or from the retrieval of an instance (a specific memory for prior performance of that specific task). In Logan's theory, the algorithm and the instances compete. Similarly, some models of categorization (Nosofsky, Palmeri, & McKinley, 1994; Erickson & Kruschke, 1998) hypothesize that learners both store exemplars and induce rules.

The two systems compete, but output can be based on computations from both systems (Erickson & Kruschke, 1998).

In our view, the transitive responses in our study arise from the generalization capabilities of the induced grammar, a grammar that is defined by the relative rank or strength of constraints. The alternative template responses, in contrast, reflect the memorization of templates and their application to novel words. Our findings suggest that when a template is available, it is often used.

The hypothesized template system is consonant with much psycholinguistic theory and data. Evidence for metrical templates has been provided for adults (Roelofs & Meyer, 1998) as well as for children (Gerken, 1991, 1994). The implicit priming studies of Roelofs and Meyer (1998) suggest that stress is stored in metrical frames independently of the word's segments or even CV structure. This is very similar to the proposal of metrical templates being offered by this study. Gerken (1991, 1994) accounts for the widely observed phenomenon of weak syllable omissions in child language through a model incorporating both rules for generating structures, and metrical templates representing the most frequent pattern seen in the language, to which structures can be fit. Gerken (1991) proposes that English-speaking children may develop trochaic templates due to their observation that this is the most frequent foot type in the language. Children's utterances are then frequently fit to the templates by omission of unstressed syllables of right headed feet, leaving the canonical trochaic pattern.

Although this theory fits well with that being proposed here, the question arises as to why children omit syllables rather than shifting stress to match the templates, as the adults in the current study did. One possibility is the initial low ranking of Faithfulness constraints in children's constraint systems (Smolensky, 1996a,b). Faithfulness constraints comprise a class of OT constraints that require a faithful reproduction of input structures to output structures. Smolensky proposes that the initial state of the child's constraint system is one in which these constraints are ranked below structural

constraints such as those used in our experiments. If the metrical templates proposed by Gerken (1991, 1994) can be thought of as structural constraints requiring utterances to fit frequently heard patterns in the language, syllable omissions would be a reasonable approach to such a fitting. Adult grammars, however, have the Faithfulness constraints ranked higher and would thus prohibit such syllable omissions, as is indeed the case (for English, at least).

While the hypothesized template system meshes with much of current psycholinguistic data and theory, it remains unclear what role an OT-like constraint system plays during online production or perception. Our finding that transitive responses are in evidence when there is no available template tells us that learners have picked up information that is, at least, associated with the OT constraints and that they have determined something of the relative strength of these information sources.

We speculate that the constraints are directly represented as weighted sources of activation that bias for particular phonological outputs. For example, any heavy syllable (or some general symbol or node representing "heaviness") would be associated with a very strong connection weight between it and the assignment of stress. Similarly, the assignment of stress to any syllable would compete, for example, though mutual inhibitory connections, with that to any adjacent syllable, thus creating a bias that prevents clash. The learning process would set the weights of these connections. Direct conflicts between CLASH and HEAVY would be particularly informative, of course. But the weights would also be influenced by examples that do not put these into direct conflict. For example, learning that CLASH is more important than some constraint, *X*, could strengthen the hypothesized inhibitory connections that are associated with preventing stress clash. Other examples showing that *X* is more important than HEAVY would then lead to the weakening of the weights associating heavy syllables with stress. Thus, even if the examples do not directly place two constraints into conflict, the weight modifications support the learning of their ranking, or transitivity. Beyond these spec-

ulations, though, we cannot make any claims about the role of the grammar in online processing. In fact, given the strong effect of templates found in our study and in other studies, we believe that the templatic system is by far the major determinant of behavior in, for example, language production by normal speakers of languages like English.

To see the effect of the grammar, one would have to look at situations in which templates are relatively unavailable, as in Experiments 2 and 3. However, our use of patently artificial languages is, of course, less than ideal. Extending our conclusions to natural languages and learning situations would be a challenge. It would require finding languages, people, or situations in which novel, but grammatical, stress patterns need to be correctly processed. Clearly, these are likely to be found early in the acquisition process (either native or nonnative acquisition). Also, one may be able to test the assumption using forms in languages with extremely productive morphologies, by devising situations in which novel, or at least rare, outputs are required.

## CONCLUSIONS

OT provides a unique linguistic framework that has many potential applications to different aspects of language, including the study of language learning and processing (e.g., Berent, Everett, & Shimron, 1998; Hammond, 1995, Levelt, 1994), linguistic analyses (e.g., Cole & Kisseberth, 1994; Hammond, 1997; McCarthy & Prince, 1993), and clinical applications (Bernhardt & Stemberger, 1998; Stemberger & Bernhardt, 1997). The results of our experiments add to these efforts by testing a key assumption of OT and by suggesting how such a grammar might coexist with other psycholinguistic mechanisms.

## APPENDIX

This appendix details the metrical analysis of the LA stress system in terms of a ranked set of constraints on the structure and location of metrical feet within the prosodic word. The stress system of LB is subject to an analysis parallel to the one presented here for LA, with substitution of a few mirror-image metrical constraints. The four basic properties of the LA stress system are 1. The penultimate

syllable of the word is stressed. 2. Stress is placed on every odd-numbered syllable counting from the beginning of the word, excluding the final syllable. 3. Heavy syllables are stressed. 4. Adjacent syllables may not both be stressed (no stress clash).

The stress system can be characterized in metrical terms as a trochaic, quantity-sensitive system of rhythmic stress. The metrical structures that underlie LA stress are illustrated in the following example of a word from the LA training set, with stress on the first and fourth light syllables. Metrical feet are parsed with parentheses notation.

(i) ( $\acute{\sigma}\sigma$ )  $\sigma$  ( $\acute{\sigma}\sigma$ )

The basic metrical unit in LA is a left-headed (trochaic) foot, ( $\acute{\sigma}\sigma$ ) or ( $\acute{\sigma}_H$ ), which results from the following constraints governing metrical parsing, adapted from McCarthy and Prince (1995).

(ii) PARSE(mora): Every mora of the input must be parsed in the prosodic structure of the output.

(iii) FOOT-FORM(trochaic): Feet are left-headed.

(iv) BINARITY(moraic): Feet must be binary (parsing moras).

The regular occurrence of stress on the penultimate syllable in LA is accomplished through the ALIGN-PrWd(right) constraint (v).

(v) ALIGN-PrWd(right): Every Prosodic Word is aligned at its right edge with the right edge of a metrical foot.

The sequence of alternating stresses on odd-numbered syllables starting from the left edge of a word is the consequence of another alignment constraint, ALIGN-FOOT(left) (vi), this time requiring that every foot in the prosodic word be aligned at the left edge of the word.

(vi) ALIGN-FOOT(left): Every foot is aligned at its left edge with the left edge of the Prosodic Word.

The effect of ALIGN-FOOT(left), due to its gradient evaluation (see McCarthy and Prince, 1995), is to anchor a series of alternating stresses at the left edge of the word, with no unparsed syllables separating the feet.

LA words with an odd number of light syllables introduce a conflict between ALIGN-PrWd(right) and ALIGN-FOOT(left). The source of the conflict lies in the fact that given an odd number of moras it is not possible to parse every mora into a binary foot. Barring deletion of a mora, which never occurs in LA, all candidates will either leave one mora (and thus one syllable) unparsed or parse a degenerate, unary foot in such words. Like most natural stress systems, LA does not tolerate unary feet (a fact reflected in the total absence of adjacent stressed syllables), which can be attributed to the highly ranked BINARITY constraint. Among the candidates with only binary feet, the location of the unparsed mora is determined by the ranking of the two alignment constraints, as shown in Table A1.

In LA the placement of stress on the penultimate syllable

TABLE A1  
BINARITY, ALIGN-PrWd(right) >> ALIGN-FOOT(left)

$\sigma\sigma\sigma\sigma$	$\mathbb{F}(\acute{\sigma}\sigma)\sigma(\acute{\sigma}\sigma)$	$(\acute{\sigma}\sigma)(\acute{\sigma}\sigma)\sigma$	$(\acute{\sigma}\sigma)(\acute{\sigma})(\acute{\sigma}\sigma)$	$(\acute{\sigma}\sigma)(\acute{\sigma}\sigma)(\acute{\sigma})$
BINARITY			*!	*!
ALIGN-PrWd(right)		*!		
ALIGN-FOOT(left)	!			

takes priority over the left alignment of the sequence of alternating stresses, modeled by ranking ALIGN-PrWd(right) over ALIGN-FOOT(left). This ranking underlies the ranking of PENULT >> LEFT-ALTERNATING described in earlier sections of this paper.

Most of the word types of LA contain heavy syllables, and their analysis in terms of constraints on metrical structure will require the introduction of several additional constraints. First, a few remarks about the nature of heavy and light syllables in LA and LB are in order. Both LA and LB distinguish the “heavy” CVVC syllable [to:n] from the “light” CVV syllables [do:], [re:], etc. Both syllable types contain long vowels, at least when viewed from the perspective of English, which contrasts the vowels of the LA/LB syllabary with a set of short, nondiphthongal vowels. Contrastively long vowels are analyzed in phonological theory as bimoraic (bearing two weight units), and thus it may be argued that the weight distinction in LA and LB is really a distinction between heavy (bimoraic) CVV syllables and the superheavy (trimoraic) CVVC syllable. While such a distinction is typologically rare, it is not unattested. The stress system of Hindi, for example, makes a three-way weight distinction: CVVC, CVCC > CVV, CVC > CV, with stress assigned to the rightmost, heaviest syllable, provided it is nonfinal (Hayes 1995:162–167). Another perspective on the LA/LB weight distinction is possible if one

takes into account the absence of a vowel weight contrast in the LA/LB syllabary. In the absence of a contrast between long and short vowels, it is consistent with standard phonological theory to assign a single mora to the (phonetically long) vowel in all of the LA/LB syllables. Then, only the [to:n] syllable would qualify as bimoraic, on the basis of its coda consonant, which makes the weight contrast in LA and LB of the standard variety. We adopt the latter perspective here and find support for this view from the tendency of some subjects to reduce the vowel length and quality of the light syllables in unstressed positions, thereby producing a true CV syllable.

The quantity-sensitive nature of LA stress is exemplified by words of the type  $\acute{\sigma}\sigma\acute{\sigma}_H$ , among others. Without distinguishing heavy syllables from light, the ranking of ALIGN-PrWd(right) >> ALIGN-FOOT(left) would locate stress on the penultimate syllable, yielding the incorrect structure  $\sigma(\acute{\sigma}\sigma_H)$ . The LA stress pattern for this sequence requires a metrical structure like  $(\acute{\sigma}\sigma)(\acute{\sigma}_H)$ . The incorrect  $\sigma(\acute{\sigma}\sigma_H)$  structure can be eliminated on the basis of two constraints. BINARITY rejects the  $(\acute{\sigma}\sigma_H)$  foot on the basis that it parses three moras rather than two (assuming that heavy syllables are bimoraic). The PEAK-PROMINENCE constraint (Prince & Smolensky, 1993:39 OT) also rejects the  $(\acute{\sigma}\sigma_H)$  foot on the basis that the head element, a light syllable, has a lower weight prominence than the non-head syllable. The structure that correctly determines the LA stress pattern,  $(\acute{\sigma}\sigma)(\acute{\sigma}_H)$ , satisfies both BINARITY and PEAK-PROMINENCE. Thus, without appeal to the foot-alignment constraints the surface pattern  $\acute{\sigma}\sigma\acute{\sigma}_H$  is selected on the basis of its optimal metrical structure.

Another manifestation of quantity sensitivity in LA is seen in forms of the type  $\acute{\sigma}_H\sigma\sigma$ . Like the  $\acute{\sigma}\sigma\acute{\sigma}_H$  type, this type fails to locate stress on the penultimate syllable in favor of stressing the heavy syllable. Stressing the penult in addition to the heavy syllable would result in a stress clash

TABLE A2  
CLASH, WSP >> ALIGN-PrWd(right)

$\sigma_H\sigma\sigma$	$\mathbb{F}(\acute{\sigma}_H)\sigma\sigma^a$	$\sigma_H(\acute{\sigma}\sigma)$	$(\acute{\sigma}_H)(\acute{\sigma}\sigma)$
CLASH			*!
WSP		*!	
ALIGN-PrWd(right)	*		*

<sup>a</sup> There is another metrical structure that can give rise to the surface stress pattern of  $\acute{\sigma}_H\sigma\sigma$ , and that is  $(\acute{\sigma}_H\sigma)\sigma$ , with an initial bisyllabic foot. This foot violates the BINARITY-(moraic) constraint because it contains three moras and is excluded from the inventory of trochaic feet under the analyses of Kager (1995) and Hayes (1995). We adopt the binarity restriction here, although we note that introducing the possibility of a (HL) trochee would not substantially alter our analysis of the LA stress system.

TABLE A3  
FOOT-FORM(trochaic), BINARITY(moraic) >> PARSE

$\sigma_H\sigma\sigma$	$\mathbb{F}(\acute{\sigma}_H)\sigma\sigma$	$(\acute{\sigma}_H)\sigma(\acute{\sigma})$	$(\acute{\sigma}_H)(\sigma\acute{\sigma})$
FOOT-FORM			*!
BINARITY		*!	
PARSE	**	*	

TABLE A4  
CLASH  $\gg$  WSP, ALIGN-FOOT(left)

$\sigma_H \sigma_H \sigma \sigma$	$\not\equiv (\acute{\sigma}_H) \sigma_H \sigma (\acute{\sigma} \sigma)$	$(\acute{\sigma}_H) (\acute{\sigma}_H) \sigma (\acute{\sigma} \sigma)$	$\sigma_H (\acute{\sigma}_H) \sigma (\acute{\sigma} \sigma)$
CLASH		*!	
WSP	*		*
ALIGN-FOOT(left)	Optimal	*	*!

which, as already noted, is not a permissible structure in LA. The absolute absence of stress clash can be attributed to the CLASH constraint, which prohibits stressing adjacent syllables.

To guarantee that stress will go on the heavy syllable, and not on the penult, we appeal to the WEIGHT-TO-STRESS constraint (viii) from Prince (1990), the OT constraint that underlies our HEAVY constraint from earlier sections.

(viii) WEIGHT-TO-STRESS (WSP): Heavy syllables are prominent in foot structure (i.e., heavy syllables must be heads).

Ranking ALIGN-PrWd(right) below both CLASH and WSP results in the observed surface stress pattern for  $\sigma_H \sigma \sigma$  words, as shown in Table A2. This ranking is the basis for our earlier ranking HEAVY  $\gg$  PENULT.

The final two light syllables of  $(\acute{\sigma}_H) \sigma \sigma$  are unparsed in the above table. The highly ranked CLASH and WSP constraints succeed in eliminating candidate forms with penultimate stress. Other candidates for a  $\sigma_H \sigma \sigma$  input include forms with a final stress, which could arise only under two conditions: a final unary foot,  $(\acute{\sigma}_H) \sigma (\acute{\sigma})$ , or a final right-headed (iambic) foot,  $(\acute{\sigma}_H) (\sigma \acute{\sigma})$ . These competing candidates for an  $\sigma_H \sigma \sigma$  input are eliminated under our analysis of the LA system by ranking the constraints FOOT-FORM(trochaic) and BINARITY(moraic) above the PARSE constraint, as shown in Table A3.

Other examples illustrating the dominance of CLASH and BINARITY over PARSE include the following word types from the training set, shown here with their metrical structures:

(ix) Word types illustrating the ranking CLASH, BINARITY  $\gg$  PARSE

$(\acute{\sigma} \sigma) (\acute{\sigma}_H) \sigma \sigma$   
 $(\acute{\sigma} \sigma) (\acute{\sigma}_H) \sigma \sigma (\acute{\sigma} \sigma)$   
 $(\acute{\sigma} \sigma) (\acute{\sigma} \sigma) (\acute{\sigma}_H) \sigma \sigma$   
 $(\acute{\sigma}_H) \sigma (\acute{\sigma} \sigma) \sigma (\acute{\sigma} \sigma)$

The last example also illustrates the role of ALIGN-FOOT(left), since the medial foot is placed as far leftward as possible, subject to the prohibition on clash.

The data in the training set exhibit some further interesting consequences of the prohibition of stress clash. Words in LA of the type  $\acute{\sigma}_H \sigma_H \sigma \acute{\sigma} \sigma$  present two adjacent heavy syllables. The ban on adjacent stresses prohibits the realization of stress on both of the heavy syllables. Thus, a violation of WSP is unavoidable as long as the ban on stress clash is upheld and input syllable weight is maintained (i.e., no syllable reduction). Among the alternatives, stressing the first heavy syllable is preferred because it best satisfies ALIGN-FOOT(left), as shown in Table A4.

The details of evaluation for ALIGN-FOOT(left) are suppressed here, but include an evaluation for each foot of its distance from the left edge of the prosodic word. The optimal candidate is the one in which all the feet are as far left as possible, subject to satisfaction of more highly ranked constraints. The interested reader may consult McCarthy and Prince (1995) for a more detailed discussion of alignment theory.

A summary of the ranked constraint grammar established thus far for the LA stress system is presented in Fig. A1.

There is one word type in the LA training that remains to be discussed, and that is the type  $\sigma \sigma_H \sigma_H \sigma \sigma$ . Like the example shown in Table A4, CLASH prevents the realization of stress on the both of the adjacent heavy syllables, e.g.,  $\sigma (\acute{\sigma}_H) (\acute{\sigma}_H) \sigma (\acute{\sigma} \sigma)$ . Table A5 shows the evaluation of

TABLE A5  
BINARITY vs. PARSE

$\sigma \sigma_H \sigma_H \sigma \sigma$	$(\acute{\sigma} \sigma_H) (\acute{\sigma}_H) \sigma (\acute{\sigma} \sigma)$	$\not\equiv \sigma (\acute{\sigma}_H) \sigma_H \sigma (\acute{\sigma} \sigma)$	$\sigma \sigma_H (\acute{\sigma}_H) \sigma (\acute{\sigma} \sigma)$
BINARITY	*!		
WSP	*	*	*
PARSE	*	****	****
ALIGN-FOOT(left)	Optimal	*	*!

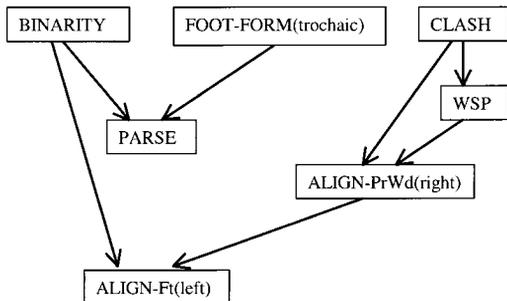


FIG. A1. Summary of Language A constraint ranking.

several clash-avoiding parsed candidates for a  $\sigma\sigma_H\sigma_H\sigma\sigma$  input.

The second and third candidates fare equally on the top three constraints in the table and differ only in the left alignment of their feet. The first candidate presents the surface stress pattern, but under this evaluation it should be eliminated because it violates the top-ranked BINARITY constraint with its parse of an initial trimoraic foot,  $(\acute{\sigma}\sigma_H)$ . Based on the table above, the optimal candidate for this input should be the second one. The surface stress pattern for this input,  $\acute{\sigma}\sigma_H\acute{\sigma}_H\acute{\sigma}\sigma$ , can, however, be generated without a BINARITY violation once we allow for the possibility that the  $(\acute{\sigma}\sigma_H)$  foot, itself a violation of PEAK-PROMINENCE, can be improved by reducing the weight of the heavy syllable. Trochaic shortening is the term used for the reduction of a unbalanced trochee to a balanced  $(\acute{\sigma}\sigma)$  trochee. The surface stress pattern of  $\acute{\sigma}\sigma_H\acute{\sigma}_H\acute{\sigma}\sigma$  in LA is consistent with a trochaic shortening analysis, as shown in Table A6.

Shortening incurs a violation of PARSE(mora), since in shortening the heavy syllable one of its underlying moras fails to be parsed. For this reason the first candidate has one more PARSE violation in this table than it shows in Table A5. The second and third candidates are eliminated in favor of the first by their greater number of PARSE or ALIGN-FOOT(left) violations. Since these two constraints are unranked with respect to one another, there is an ambiguity as to which constraint marks the fatal violation, which is noted in the table by enclosing fatal marks for both constraints in parentheses.

We find evidence for the Trochaic Shortening analysis of

TABLE A7

Test Item:  $\sigma\sigma_H\sigma\sigma$

$\sigma\sigma_H\sigma\sigma$	$\text{[}\acute{\sigma}\text{]}\sigma(\acute{\sigma}_H)\sigma(\acute{\sigma}\sigma)$	$(\acute{\sigma}\sigma_H)\sigma(\acute{\sigma}\sigma)$	$(\acute{\sigma}\sigma)\sigma(\acute{\sigma}\sigma)$
BINARITY		*!	
WSP		*	*!
PARSE	**	*	**
ALIGN-FOOT(left)	*	Optimal	Optimal

$\sigma\sigma_H\sigma_H\sigma\sigma$  input in the observation that many subjects produced a reduced vowel quality ([ə]) for the unstressed heavy syllable in words such as DO [tə] TON re MI fa.

We turn now to the metrical analysis of the two test items in LA. These input forms can be evaluated by the set of constraints under the ranking established on the basis of the training items. The three best competitors are shown in Table A7 for the first test item,  $\sigma\sigma_H\sigma\sigma$ . Since these three candidates equally satisfy the FOOT-FORM, CLASH, and ALIGN-PrWd(right) constraints, these constraints are not shown in the table.

Candidate 2 fails on BINARITY, known to be an undominated constraint in LA. Candidates 1 and 3 tie on PARSE, though for different reasons: candidate 1 leaves two light syllables out of the metrical parse, while candidate 3 leaves one light syllable unparsed and shortens a heavy syllable to a light parsed syllable. The decision between 1 and 3 boils down to the ranking between WSP and ALIGN-FOOT(left). These two constraints do not interact directly in the analysis of any of the training items, and this is where transitivity comes into play. The training items establish the two independent rankings  $WSP \gg ALIGN-PrWd(right)$  and  $ALIGN-PrWd(right) \gg ALIGN-FOOT(left)$ , the transitive closure of which yields the ranking  $WSP \gg ALIGN-FOOT(left)$ . This transitive ranking is the grounds on which candidate 3, the trochaic shortening candidate, is eliminated in favor of candidate 1.

A similar situation obtains in the evaluation of the second LA test item,  $\sigma\sigma\sigma_H\sigma\sigma$ . In fact, the only difference between the two test items is in the presence of two additional light syllables at the beginning of  $\sigma\sigma\sigma_H\sigma\sigma$ . These two light syllables form a perfect trochaic foot,  $(\acute{\sigma}\sigma)$ , and since the remainder of the string is identical to the first test item, the

TABLE A6

Trochaic Shortening

$\sigma\sigma_H\sigma_H\sigma\sigma$	$\text{[}\acute{\sigma}\text{]}\sigma(\acute{\sigma}_H)\sigma(\acute{\sigma}\sigma)$	$\sigma(\acute{\sigma}_H)\sigma_H\sigma(\acute{\sigma}\sigma)$	$\sigma\sigma_H(\acute{\sigma}_H)\sigma(\acute{\sigma}\sigma)$
BINARITY			
WSP	*	*	*
PARSE	**	***(!)*	***(!)*
ALIGN-FOOT(left)	Optimal	*(!)	*(!)

TABLE A8

Test Item:  $\sigma\sigma\sigma_{\text{H}}\sigma\sigma$ 

$\sigma\sigma\sigma_{\text{H}}\sigma\sigma$	$\sigma(\acute{\sigma})\sigma(\acute{\sigma}_{\text{H}})\sigma(\acute{\sigma}\sigma)$	$(\acute{\sigma}\sigma)(\acute{\sigma}\sigma_{\text{H}})\sigma(\acute{\sigma}\sigma)$	$(\acute{\sigma}\sigma)(\acute{\sigma}\sigma)\sigma(\acute{\sigma}\sigma)$
BINARITY		*!	
WSP		*	*!
PARSE	**	*	**
ALIGN-FOOT(left)	*	Optimal	Optimal

evaluation of the candidate forms is parallel to the evaluation shown in Table A7. Table A8 spells out the details of evaluation for the test item  $\sigma\sigma\sigma_{\text{H}}\sigma\sigma$  and illustrates once again the role of the transitive ranking  $\text{WSP} \gg \text{ALIGN-FOOT(left)}$  in selecting the optimal candidate, candidate 1.

## REFERENCES

- Archangeli, D., & Langendoen, D. T. (Eds.) (1997). *Optimality Theory: An overview*. Malden, MA: Blackwell.
- Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **12**, 336–345.
- Berent, I., Everett, D. L., & Shimron, J. (1998). *Do phonological representations specify variables? Evidence from the obligatory contour principle*. [manuscript submitted for publication]
- Bernhardt, B. H., & Stemberger, J. P. (1998). *Handbook of phonological development from the perspective of constraint-based nonlinear phonology*. San Diego: Academic Press.
- Boysen, S. T., Berntson, G. G., Shreyer, T. A., Quigly, K. S. (1993). Processing of ordinality and transitivity by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, **107**, 208–215.
- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, **5**, 325–337.
- Christiansen, M., & Chater, N. Special issue devoted to connectionist models of language processing. *Cognitive Science*. [in press]
- Clark, H. H. (1969). Influence of language on solving three-term series problems. *Journal of Experimental Psychology*, **82**, 205–215.
- Cole, J., & Kisseberth, C. W. (1994). An optimal domains theory of harmony. *Studies in the Linguistic Sciences*, **24**, 101–114.
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, **8**, 240–247.
- Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, **100**, 589–608.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language specific comprehension strategy, *Nature*, **304**, 159–160.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable’s differing role in the segmentation of French and English, *Journal of Memory and Language*, **25**, 385–400.
- Davis, H. (1992). Transitive inference in rats (*Rattus norvegicus*). *Journal of Comparative Psychology*, **106**, 342–349.
- Erickson, M. A., Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, **127**, 107–140.
- Gerken, L. A. (1991). The metrical basis for children’s subjectless sentences. *Journal of Memory and Language*, **30**, 431–451.
- Gerken, L. A. (1994). Young children’s representation of prosodic phonology: Evidence from English-speakers’ weak syllable productions. *Journal of Memory and Language*, **33**, 19–38.
- Hammond, M. (1995). *Syllable parsing in English and French*. [unpublished manuscript]
- Hammond, M. (1997). Optimality theory and prosody. In D. Archangeli & D. T. Langendoen (Eds.), *Optimality Theory: An overview* (pp. 33–58). Malden, MA: Blackwell.
- Hayes, B. (1995). *Metrical Stress Theory*. Chicago: Univ. of Chicago Press.
- Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review*, **75**, 550–560.
- Kager, R. (1995). The metrical theory of word stress. In J. Goldsmith (Ed.), *Handbook of Phonological Theory* (pp. 367–402). Oxford and Cambridge, MA: Blackwell.
- Levelt, C. C. (1994). *On the acquisition of place*. Doctoral dissertation, Holland Institute of Generative Linguistics.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, **22**, 1–75.
- Lewicki, P., Hill, T., & Czyzewska, M. (1994). Nonconscious indirect inferences in encoding. *Journal of Experimental Psychology: General*, **123**, 257–263.
- Lipkens, R., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of the development of derived relations in

- an infant. *Journal of Experimental Child Psychology*, **56**, 201–239.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492–527.
- McCarthy, J. J., & Prince, A. S. (1995). Generalized alignment. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology* (pp. 79–153). Dordrecht: Kluwer.
- McGonigle, B. O., Chalmers, M. (1977). Are monkeys logical? *Nature*, **267**, 694–696.
- McGonigle, B. O., Chalmers, M. (1992). Monkeys are rational! *Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology*, **45B**, 189–228.
- McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18**, 1155–1172.
- McNamara, T. P. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18**, 1173–1190.
- Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, **20**, 715–726.
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, **101**, 53–79.
- Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. H. Yeni-Komshian, J. F., Kavanagh, & C. A. Ferguson (Eds.), *Child phonology, Vol. 1. Production* (pp. 93–110). New York: Academic Press.
- O'Seaghdha, P. G., & Marin, J. W. (1997). Mediated semantic-phonological priming: Calling distant relatives. *Journal of Memory and Language*, **36**, 226–252.
- Pesetsky, D. (1997). Optimality theory and syntax: Movement and pronunciation. In D. Archangeli & D. T. Langendoen (Eds.), *Optimality Theory: An overview* (pp. 134–170). Malden, MA: Blackwell.
- Piaget, J. (1967). *Six psychological studies*. New York: Random House.
- Pinker, S. (1991). Rules of language. *Science*, **253**, 530–535.
- Prince, A. S. (1990). *Quantitative consequences of rhythmic organization: Parasession on the syllable in phonetics and phonology* (pp. 355–98). Chicago Linguistic Society.
- Prince, A. S., & Smolensky, P. (1993). Optimality theory: Constraint interaction in generative grammar. Piscataway, NJ: Rutgers Univ. Cognitive Sciences Center Technical Report 2.
- Prince, A. S., & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science*, **275**, 1604–1610.
- Roelofs, A., & Meyer, A. S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **24**, 922–939.
- Rubin, D. C. (1975). Within word structure in the tip-of-the-tongue phenomenon. *Journal of Verbal Learning and Verbal Behavior*, **14**, 392–397.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing. Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
- Russell, K. (1997). Optimality Theory and morphology. In D. Archangeli & D. T. Langendoen (Eds.), *Optimality Theory: An overview* (pp. 102–133). Malden, MA: Blackwell.
- Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **18**, 1191–1210.
- Smolensky, P. (1996a). The initial state and 'richness of the base' in Optimality Theory. Technical Report JHU-CogSci-96-4. Baltimore: Johns Hopkins Univ. Department of Cognitive Science.
- Smolensky, P. (1996b). On the comprehension/production dilemma in child language. *Linguistic Inquiry*, **27**, 720–731.
- Speas, M. (1997). Optimality Theory and syntax: Null pronouns and control. In D. Archangeli & D. T. Langendoen (Eds.), *Optimality Theory: An overview* (pp. 171–199). Malden, MA: Blackwell.
- Stark, R. E. (1980). Stages of speech development in the first year of life. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology, Vol. 1, Production* (pp. 73–90). New York: Academic Press.
- Steirn, J. N., Weaver, J. E., & Zentall, T. R. (1995). Transitive inference in pigeons: Simplified procedures and a test of value transfer theory. *Animal Learning and Behavior*, **23**, 76–82.
- Stemberger, J. P., & Bernhardt, B. H. (1997). Optimality Theory. In M. J. Ball & R. D. Kent (Eds.), *The new phonologies: Developments in clinical linguistics* (pp. 211–245). San Diego: Singular.
- Tesar, B., & Smolensky, P. (1993). The learnability of Optimality Theory: An algorithm and some basic complexity results. Boulder, CO: Technical Report, Department of Computer Science, Univ. of Colorado.
- Trabasso, T. (1977). The role of memory as a system in making transitive inferences? In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition*. Hillsdale, NJ: Erlbaum.
- Treichler, F. R., & Tilburg, D. V. (1996). Concurrent conditional discrimination tests of transitive inference by macaque monkeys: List linking. *Journal of Experimental Psychology: Animal Behavior Processes*, **22**, 105–117.
- Vihman, M. M. (1996). *Phonological development. The origins of language in the child*. Cambridge, MA: Blackwell.
- Weaver, J. E., Steirn, J. N., & Zentall, T. R. (1997). Transitive inference in pigeons: Control for differential value transfer. *Psychonomic Bulletin and Review*, **4**, 113–117.

(Received February 17, 1999)

(Revision received September 1, 1999)