The double identity of doubling: Evidence for the phonology–morphology split

Iris Berent
Northeastern University

Outi Bat-El
Tel-Aviv University

Vered Vaknin-Nusbaum
Western Galilee College

Address for correspondence:
Iris Berent
Professor
Department of Psychology
Northeastern University
i.berent@neu.edu
125 Nightingale
360 Huntington Ave
Boston MA 02115
Phone (617) 373 4033
Fax: (617) 373-8714

Abstract

Duality of patterning, is, by hypothesis, a universal design feature of language. Every language constructs words from meaningful units (morphemes), which, in turn, are comprised of meaningless phonological elements (e.g., segments, syllables). But whether the language faculty does, in fact, include a separate morphological level, distinct from the phonology, is a matter of controversy. To elucidate the role of morphology, here we ask whether morphological forms are constrained by putatively universal combinatorial principles, distinct from those applying to phonological patterns. Our research exploits the structural ambiguity of doubling. Doubling (e.g., *trafraf*) is open to two competing interpretations — as either a purely phonological form, or as a complex morphological structure that is systematically linked to meaning (e.g., *trafraf* is the diminutive of *traf*). Our experiments show that responses to doubling (*trafraf*) shift radically, depending on its level of analysis. Viewed as a meaningless phonological form, doubling is dispreferred irrespective of its kind (i.e., *trafraf* is as bad as *traftaf*, even though the latter violates a morphological constraint on contiguity). But once doubling is systematically linked to meaning (i.e., as a morphological structure), the doubling dislike shifts into a reliable preference, and an additional constraint on its contiguity arises (i.e., *trafraf* > *traftaf*). Remarkably, the dissociation between morphological and phonological doubling emerges regardless of whether morphological reduplication is abundant in participants' language (in Hebrew) or relatively rare (in English). These results suggest the existence of distinct linguistic constraints that preferentially target the morphological vs. phonological levels. We discuss various explanations for the origins of these restrictions.

1. Introduction

Duality of patterning, is, by hypothesis, a universal design feature of language (Hockett, 1960). Words are constructed from meaningful units (morphemes), which, in turn, are comprised of meaningless phonological elements (e.g., segments, syllables). A large body of psycholinguistic research has gauged the psychological reality of morphology by examining whether complex words (e.g., *liked*) are decomposed onto morphological constituents (e.g., the base *like*; Feldman & Bentin, 1994; Frost, Deutsch, & Forster, 2000; Marslen-Wilson, 2005; Rastle, Davis, & New, 2004; Taft & Forster, 1975). But since related words (e.g., *like-liked*) usually share phonological, semantic and orthographic features, the dissociation of morphemes from their correlates presents a formidable challenge.

The present research approaches the problem from a different perspective. Rather than searching for the elusive building blocks of morphology, here, we seek to elucidate the principles that govern their combinations. We ask whether patterning at the morphological level is preferentially subject to putatively universal linguistic restrictions, distinct from those applying to meaningless phonological patterns. To the extent that the constraints on morphology and phonology are distinct, then these two levels of representation must be separate. Our research examines whether such distinct sets of constraints exist, and whether their knowledge requires extensive linguistic experience.

To address these questions, we exploit the structural ambiguity in the interpretation of doubling. Doubling (e.g., *banana, panana*) is amenable to two competing interpretations — as either purely phonological forms (e.g., English: *banana*), or as a complex morphological structure, where doubling indicates systematic links between form and meaning (e.g., Manam: *pana* 'chase'→ *panana* 'run'; Lichtenberk, 1983). Moreover, doubling at the phonological and morphological levels is subject to distinct sets of constraints that are putatively universal.

Our experiments demonstrate that these constraints guide language processing. We show that responses to doubling (e.g., *trafraf*) shift radically, depending on the level of its analysis — as either a meaningless phonological pattern, or a complex morphological structure (e.g., *trafraf* is the diminutive of *traf*). Given that the input that elicits these different responses is invariant, the shift in response must reflect distinct principles that operate at the phonological and morphological levels. Furthermore, our experiments show that participants exhibit knowledge of these principles despite only limited experience with morphological doubling in their own language. These findings suggest that morphology is an autonomous component of the language system, distinct from the phonology. We discuss various functional explanations for the origins of the restrictions on each level.

2. The double identity of doubling

Practically every known language includes restrictions that specifically target doubling (Suzuki, 1998; Walter, 2007). The nature of those restrictions, however, strictly depends on the structural parse of those doubling elements — at the phonology or the morphology (Inkelas, 2008).

At the morphological level, doubling is the product of ***reduplication***— a productive process that generates complex morphological forms by copying a base, either fully or partially (Marantz, 1982; McCarthy & Prince, 1995a; Wilbur, 1973). For example, the Hebrew base *katan* 'small' gives rise to *ktantan* 'smallish' — a complex reduplicative form that shares with the base both form and meaning. In the Hebrew case, the -*tan* element (the reduplicant) is clearly drawn from *katan*, so *ktantan* is derived from *katan* (cf. *xalaʃ* 'weak' → *xalaʃlaʃ* 'weakish').

Doubling, however, could also reflect coincidental phonological ***identity***, as in the English *Stanton* and *Trenton*. Here, *ton* (historically, a bound nominal suffix) is unrelated to *stan* — the partial repetition in *Stanton* is purely coincidental (cf. *Brighton, Houston)*. Thus, similar phonological strings may be parsed by the listener either as phonological identity (for *Stanton*) or as morphological reduplication (for *ktantan).* In what follows, we use "doubling" generically, to refer to any string that includes repeated elements; we use "identity" to refer to the representation of doubling at the level of phonology and "reduplication" to denote the encoding of doubling by the morphology. Crucially, in this analysis, these two representations (phonological identity vs. morphological reduplication) are preferentially subject to distinct constraints.

Morphological reduplication (e.g., the Hebrew *ktantan* 'smallish', from *katan* 'small') is restricted by CONTIGUITY — a constraint that governs the correspondence between the base and the reduplicant (e.g., *ktan* and *tan,* respectively). CONTIGUITY requires the reduplicant to be a contiguous linear substring of the base (McCarthy & Prince, 1995a), thus banning skipping (**ktan-kan*), insertion (**ktan-stan*), and reordering of segments (**ktan-tak*; for formal definition and the integration of the constraint LINEARITY see Appendix I). The Hebrew *ktantan*, 'smallish' obeys CONTIGUITY because *tan* is a contiguous substring of *katan*, whereas noncontiguous reduplicants (e.g., *kan* in *ktankan*), or nonlinear permutations of the base (e.g., *tak* in *ktantak*) are dispreferred.

By contrast, identical phonological forms are not required to exhibit CONTIGUITY, so *Tranton* is no worse than *Stanton*, even though *ton* is not a contiguous substring of *Tran.* In fact, identical phonological elements are often systematically avoided in phonological representations due to the OBLIGATORY CONTOUR PRINCIPLE (OCP; Leben, 1973; McCarthy, 1981).[1] And indeed, unlike reduplication, phonological identity is systematically underrepresented across languages (Suzuki, 1998; Walter, 2007),

[1] Our present analysis assumes that the OCP operates within a morpheme. We note, however, that the OCP could apply either within or across morphemes, and its effects in the two cases could differ. Furthermore, the OCP is most powerful when the elements are adjacent—the closer the distance between identical elements, the stronger their avoidance (Frisch, Pierrehumbert, & Broe, 2004; McCarthy, 1981; Rose & Walker, 2004).

including English (Berkley, 2000). Morphological reduplicants are immune to those phonological restrictions because their lexical form is free of repetition — the doublings in surface forms are merely copies (marked "c" below) of the base elements, not repeated tokens. To use an analogy, morphological reduplication is analogous to the reflection of a single individual in a mirror; phonological identity can be likened to identical twins (i.e., two tokens of a single type).

(1) Constraint violation by morphological reduplication and phonological identity

	Parse	CONTIGUITY	OCP
Morphological reduplication	{{$t_1r_2a_3f_4$} $r_{c2}a_{c3}f_{c4}$}	✓	
	{{$t_1r_2a_3f_4$} $t_{c1}a_{c3}f_{c4}$ }	*	
Phonological identity	trafraf		*
	traftaf		*

Note: The subscript *c* denotes a copy of the base segment, and the integer subscript denotes the correspondence between the copy and the base (e.g., r_{c2} is a copy of the second base segment, *r*). Note that in contiguous reduplicative forms, the integers subscripts on the copies form a contiguous substring of those denoting the base. In contrast, identical phonological elements include no copies.

Summarizing then (see 1), phonological doubling (e.g., *trafraf*) is open to conflicting interpretations — either as morphological reduplication (of *traf*) or as phonological identity, and these two parses are each subject to conflicting constraints. Identical elements are avoided in phonology (due to the OCP), whereas morphological reduplication is often encouraged, but it is required to obey CONTIGUITY.[2] We should note that our analysis is couched in the theoretical framework of Optimality Theory (McCarthy & Prince, 1995b; Prince & Smolensky, 1993/2004), where all grammatical constraints are violable, and the constraints on doubling are no exception (for violations of CONTIGUITY and the OCP, see Kager, 1999 and Berent, Everett, & Shimron, 2001, respectively).

Crucially, Optimality Theory (McCarthy & Prince, 1995b; Prince & Smolensky, 1993/2004) asserts that all grammatical constraints are universal — they are active in each and every grammar, irrespective of whether the relevant structure is present or absent in the language. This account thus predicts that CONTIGUITY and the OCP are active universally. In line with this analysis, the OCP and CONTIGUITY have been each widely documented in the formal analysis of many languages (e.g., Inkelas & Zoll, 2005; Kager, 1999; Leben, 1973; McCarthy, 1979; McCarthy & Prince, 1995b; Suzuki, 1998; Walter, 2007).[3] There is also a large experimental literature

[2] Our analysis above only lists the constraints violated at a single level (morphology vs. phonology). It is conceivable, however, that a given input could acquire distinct parallel parses at multiple levels of analysis. While at the morphological level, *trafraf*, for instance, incurs no identity violation, the same input could conceivably acquire also a secondary phonological parse where phonological identity is banned by the grammar. Our investigation asks whether the morphological and phonological levels are each associated with distinct parses. Whether secondary parses at competing levels exist remains to be seen.

[3] Although these proposals differ formally from the CONTIGUITY constraint, they nonetheless echo a similar general principle inasmuch as the reduplicant material is required to exhibit identity (Wilbur, 1973) or coupling with the base (Zuraw, 2002), preserve the left-to-right ordering of the base

demonstrating that speakers productively extend identity restrictions to novel forms (e.g., Berent & Shimron, 1997; Berent et al., 2001; Berkley, 1994; Boll-Avetisyan & Kager, 2014; Buckley, 1997; Frisch & Zawaydeh, 2001; Kawahara, Ono, & Sudo, 2006). However, no previous experimental research has examined whether people productively obey CONTIGUITY. Accordingly, it is unclear whether the constraints on doubling are distinct at the morphological and phonological levels. It is also unknown whether sensitivity to these constraints can emerge in the absence of extensive linguistic experience with morphological reduplication.

Our investigation takes the first step to address these questions. To gauge sensitivity to CONTIGUITY, we examine whether this constraint applies selectively at the morphological (but not phonological) level. We first ask whether speakers productively enforce CONTIGUITY when morphological reduplication is prevalent in their native language; Hebrew presents a case in point (Experiment 1). Having shown that CONTIGUITY is productive when reduplication is abundant, we next move to ask whether people spontaneously converge on the same reduplicative parse when provided with more limited experience with morphological reduplication; to this end, we present similar forms to speakers of English (Experiment 2). Results suggest that, absent an overt morphological context, English speakers parse ambiguous forms as exponents of phonological identity: they dislike all forms of identity (e.g., *trafraf, traftaf*), with no preference for contiguous stimuli (e.g., *trafraf*). Remarkably, once the morphological link to the base is established, the preference for contiguous forms emerges also for English speakers (Experiment 3).

3. Hebrew speakers: Phonological forms (Experiment 1a-b)

Modern Hebrew uses reduplication quite frequently (see examples in (2)). While some forms of reduplication have a semantic function (e.g., in forming diminutives and in marking durative/repetitive meanings, Bolozky, 1999; Ussishkin, 1999), others lack systematic semantic links (see Bat-El 2006), either in form ((2d); orphan forms), or in meaning (2c).[4] The Hebrew data in (2) below show that Hebrew reduplicated forms do not violate CONTIGUITY. Our question here is whether CONTIGUITY is a productive grammatical constraint.

elements (Marantz, 1982, principle D), maintain precedence in the linear order of segments (Raimy, 2012) or avoid "line crossing" (Frampton, 2009; McCarthy, 1979).

[4] Since orphan forms (e.g., *ʃravrav* 'plumber') follow the morphological structure of attested reduplicative forms (e.g., *klavlav,* 'little dog'), Bat-El (2006) analyzed both as reduplicative, but whether Hebrew speakers do in fact parse orphan forms as morphological reduplication or phonological identity is unknown—our present experiments test this question.

(2) Reduplicative forms in Modern Hebrew

		Reduplicative		Related form	
a.	Diminutives	*klavlav*	'little dog'	*kelev*	'dog'
		ktantan	'smallish'	*katan*	'small'
b.	Durative/repetitive	*dimem*	'bled'	*dam*	'blood'
		kidrer	'dribbled'	*kadur*	'ball'
c.	Others	*xagag*	'celebrated'	*xag*	'holiday'
		ʃagrir	'ambassador'	*ʃiger*	'dispatched'
d.	Orphan forms	*parpar*	'butterfly'	(not related to *par* 'bull')	
		ʃravrav	'plumber'	(not related to *ʃarav* 'heat')	
		yaʃiʃ	'old man'	(*yaʃ* does not exist)	

To address this question, we compared responses to three types of novel forms: (a) reduplicative forms obeying CONTIGUITY (e.g., *trafraf*); (b) reduplicative forms violating CONTIGUITY (e.g., *traftaf* or *trafrat*);[5] and (c) non-reduplicative controls (e.g., *trafkam*). Participants were asked to rate these triplets either relative to each other (*relative rating*) or in a randomized list (*absolute rating*).

Since morphological reduplication is productive in Hebrew, then by default, participants should attempt to apply it to the ambiguous input, and evaluate it for CONTIGUITY violations. Accordingly, *trafraf* (which abides by CONTIGUITY) should be preferred to the noncontiguous *traftaf*. Furthermore, since in *trafraf*, the base is free of doubling (e.g., $\{\{t_1r_2a_3f_4\}r_{c2}a_{c3}f_{c4}\}$), it includes fewer phonological elements than the control *trafkam.* Accordingly, we expect Hebrew speakers to prefer the (CONTIGUITY obeying) reduplicated forms (e.g., *trfafraf*) to their non-reduplicative counterparts (e.g., *trafkam*).

3.1. Methods

3.1.1. ***Participants:*** Two groups of participants took part in the two rating procedures. Each group consisted of thirty native Hebrew speakers, students at Western Galilee College, Israel.

3.1.2. ***Materials:*** The experimental materials included 40 novel word triplets, a total of 120 items (Appendix II). Each triplet consisted of three matched forms: contiguous reduplicatives, noncontiguous reduplicatives and controls. Contiguous reduplicatives exhibited a reduplicant that is a contiguous linear string in the base (e.g., *trafraf*); Noncontiguous reduplicatives violated the contiguity requirement, either because their segments were nonadjacent in the base (e.g., *traftaf*), or because they violated the linear order of the base's segments (e.g., *trafrat*).[6] Finally,

[5] Unlike contiguous forms, non-contiguous forms are not a natural class in linguistic theory, as there are many ways to violate CONTIGUITY. We chose *traftaf* (i.e., a "no-skipping" violation) and *trafrat* (a violation of "no reordering of segments") for our "non-contiguous" type because these forms match the contiguous member (*trafraf*) for length (unlike violations that insert segments, such as *trafraft*).

[6] Upon further inspection, we noted that six of the noncontiguous items (*dvandav, zgavzag, ʃgadʃag, pdanfad, gzavgaz, pkanpak*) exhibited an $\{XYZ\}X_cY_c$ structure (instead of the intended $\{XYZ\}X_cY_c$ form). These items were thus excluded from all analyses.

the control condition paired the base with an unrelated syllable (e.g., *trafkam*), which does not form a suffix in the language.

The two sub-types of non-contiguous items (e.g., *traftaf* vs. *trafrat*) were originally introduced in order to shed light on the precise definition of contiguity (i.e., whether contiguous reduplicants must also preserve the linear order of the elements in the base; see Appendix I). An analysis of the results using Wilcoxon matched pairs test found that responses to the two types of non-contiguous forms (e.g., *traftaf* or *trafrat*) were virtually identical in each of the two rating experiments (all $p<.31$ by participants and items). For the sake of simplicity, we thus proceeded to conduct all analyses while ignoring this factor. Accordingly, the means for the non-reduplicative forms are always reported while collapsing across the two subtypes.

All items and their bases were novel Hebrew words that were phonotactically legal in Hebrew.

3.1.3. ***Procedures:*** Each participant took part in one of two rating procedures, eliciting acceptability ratings of the stimuli as Hebrew words. The *relative rating procedure* presented each triplet as a single group (with order counterbalanced), and participants were asked to rate its members relative to each other on a 1-3 scale (1=best, 2=intermediate, 3=worst). In the *absolute rating procedure*, all 120 items were mixed in a single list (with order randomized), and participants were asked to rate each item on its own, using a 1-5 scale (1=worst, 5=best).

We chose to include both procedures because these two tasks differ in the extent to which they explicitly require participant to attend to the internal structure of the stimulus, and our past research observed some differences in their outcomes (Berent et al., 2001; Everett & Berent, 1999). The *relative rating* contrasts matched items (e.g., *trafraf* vs. *traftaf*), so it may elicit greater sensitivity to the internal reduplicative structure of the stem. By contrast, the *absolute* rating task can be informed by attending to any aspect of the individual stimulus. Accordingly, convergence across the two tasks, regardless of whether attention to reduplication is encouraged (in *relative* rating) or not (in *absolute rating*), would suggest that people represent CONTIGUITY automatically, even when attention to reduplication is not promoted by the task.

In each task, all items were presented in print (with all vowels specified using diacritics).[7] For viewing convenience, we invariably report our findings such that preference is expressed by higher numerical values. For the relative rating, we subtracted the means from the constant 4. The wording of the instructions for the relative rating experiment are provided below (in English translation); the absolute rating instructions were identical, except that participants were asked to rate each word in isolation on a 1-5 scale (1=very bad; 5=excellent).

[7] The Hebrew writing system is mostly consonantal; vowels are usually specified by diacritics. Although most texts do not use the vowel diacritics, most literate Hebrew speakers are adept at decoding the diacritic vowel system as this system is widely used in children's books, religious texts and poetry.

"*In this experiment, we created a bunch of words — these words are not real Hebrew words, but in our opinion, some of them sound quite Hebrew-like. We would like to know your opinion: which of these words sound like Hebrew?*

To find out, we arranged the words in triplets. We ask you to read the words in each triplet aloud, pronouncing them exactly as they are printed. Then, please indicate which of the three words is the best, which one is the worst and which one is intermediate. We ask you to only pay attention to the sound — do not try to associate the words with the meaning of any word or word parts — only the sound matters. If the word sounds the best, please indicate 1; if it's intermediate, please indicate 2; if it's the worst, please indicate 3. Thank you very much!"

3.2. Results

Figure 1 plots the effect of reduplication on acceptability in the relative and absolute rating tasks. An inspection of the means suggests that Hebrew speakers favored contiguous reduplication (*trafraf*) to either non-contiguous reduplication (*traftaf*) or no reduplication (*trafkam*), and this preference obtained irrespective of the rating procedure.[8]

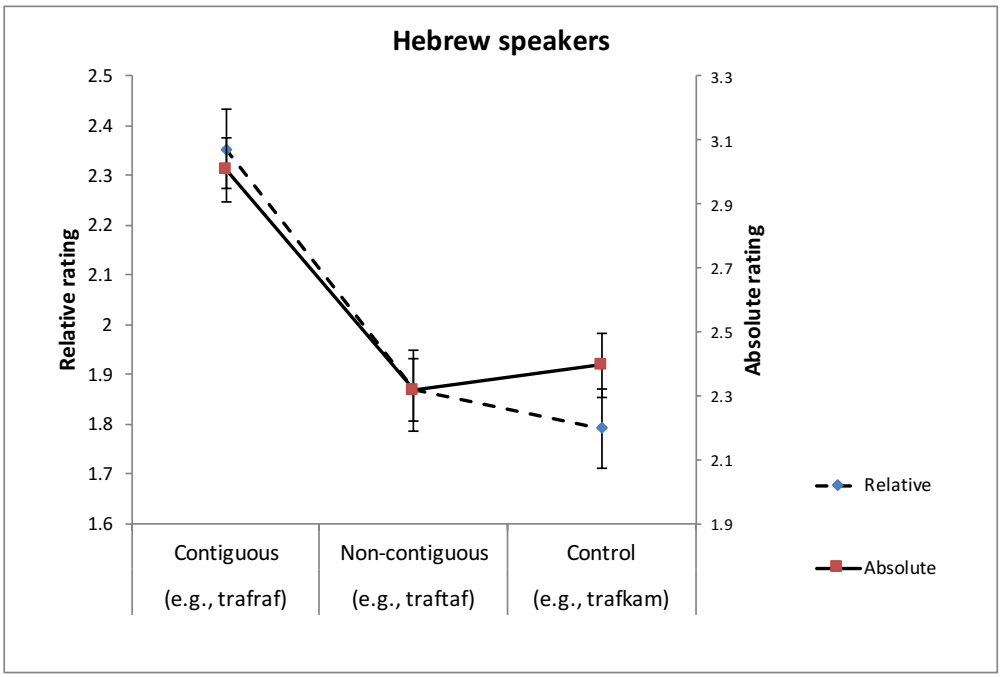

Figure 1. Acceptability ratings in Experiment 1. Error bars are 95% confidence intervals for the difference between the means.

[8] In Figure 1 and all subsequent figures, error bars reflect 95% confidence intervals for the difference between the means, created using the error term from an ANOVA conducted over participants' means.

These conclusions are also borne out by the statistical analyses (Table 1). To determine whether people were sensitive to the reduplication type, we first compared the acceptability of the three types of items by means of a Friedman nonparametric ANOVA, conducted using both participants and items as random variables. The effect of type was significant for both the relative ($X^2_1(2)=29.87$, $p<.0001$; $X^2_2(2)= 27.00$, $p<.0001$) and absolute ($X^2_1(2)=20.66$, $p<.0001$; $X^2_2(2)=31.60$, $p<.0001$) rating procedures. Subsequent nonparametric tests of the contrasts between the means, using the Wilcoxon Matched Pairs Test, confirmed that participants rated contiguous reduplicants higher than either non-contiguous or non-reduplicant types, which, in turn did not differ.

Table 1. Pairwise statistical contrasts in Experiment 1

Procedure	Contrast	Participants		Items	
		Z	p	Z	p
Relative rating	trafraf vs. traftaf	4.42	0.0001	3.92	0.0001
	trafraf vs. trafmak	4.66	0.0001	3.85	0.0001
	traftaf vs. trafmak	1.22	0.22	1.23	0.21
Absolute rating	trafraf vs. traftaf	5.00	0.0003	5.51	0.0001
	trafraf vs. trafmak	19.00	0.0012	5.36	0.0001
	traftaf vs. trafmak	0.47	0.65	1.44	0.15

3.3. Discussion

The results of Experiment 1 demonstrate that Hebrew speakers favor novel contiguous reduplicative forms (e.g., *trafraf*) to either noncontiguous reduplicatives (e.g., *traftaf*) or nonredupliative controls (e.g., *trafkam*), which, in turn did not differ.[9] This finding is consistent with the possibility that speakers parse inputs such as *trafraf* as exponents of reduplication, and they require that reduplicative forms must form a contiguous linear substring of the base. Accordingly, contiguous items like *trafraf* are preferred.

Before accepting this conclusion, however, we must first consider an alternative explanation for the results. This alternative account asserts that the preference for contiguous items (e.g., *trafraf*) is due not to their relation to the base (e.g., to *traf*) but rather to the phonological properties of the reduplicative form itself. One version of this hypothesis states that Hebrew speakers prefer *trafraf* because the onset of its second syllable (e.g., ***r****af*) is more frequent in the Hebrew lexicon (relative to ***t****af*, in *traftaf*). To address this possibility, we evaluated the frequency of our reduplicative onsets in Hebrew nouns (based on the statistical analysis conducted by Klein (2015) on Becker and Bolozky's 2006 database). Contrary to the frequency account, however, the frequency of the consonant-vowel sequence in congruent reduplicants (e.g., *ra* in *trafraf*, M=291, SD=182) did not differ reliably

[9] How speakers encode noncontiguous forms like *traftaf* is less clear from these findings. One possibility is that the violation of CONTIGUITY prevented speakers from encoding reduplication altogether. Alternatively, *traftaf* might be encoded as a defective reduplicative form, but the violation of CONTIGUITY might offset any advantages of reduplicative forms over controls. Further research is required to adjudicate between these possibilities.

from the noncongruent reduplicants (e.g., *ta* in *traftaf*, M=254, SD=195; t(38)=1.30, p>.20, n.s.). Accordingly, it is unlikely that the preference for *trafraf* is due to statistical properties alone.

It is also unlikely that the CONTIGUITY preference is due to a structural phonological preference. One possibility is that the contiguous item is preferred not because of CONTIGUITY per se but rather due to its syllabification. Indeed, sequences like *trafraf* and *traftaf* require that participants parse the internal consonant cluster. One could assume that in congruent forms like *trafraf*, the cluster can form an onset (e.g., *tra.fraf*), whereas in the non-contiguous alternative, the cluster must be parsed as coda-onset sequence (e.g., *traf.taf*). Accordingly, CONTIGUITY might be confounded with syllabification. However, this is not the case for Hebrew phonology. Unlike English, Hebrew allows for a rich set of complex onsets, including not only obstruent-sonorant combinations (e.g., *tris* 'shutter') but also obstruent-obstruent sequences (e.g., *ʃ vil* 'trail', *bgida* 'betrayal'), so contiguous and noncontiguous forms are both amenable to the same syllabic parses—either as complex onsets (e.g., *tra.fraf, tra.ftaf*) or as coda-onset sequence (e.g., *traf.raf, traf.taf*). Consequently, the preference for contiguous reduplicants cannot be explained by syllabification.

It is also unlikely that the CONTIGUITY parse reflects a structural preference for sonorant relative to obstruent onsets (*raf* vs. *taf*). In fact, sonorant onsets are generally dispreferred across languages (Clements, 1990), a trend evident experimentally in adults (e.g., Stemberger & Treiman, 1986) and children (e.g., Ohala, 1999). Moreover, our results directly speak against this explanation. If the dispreference of noncontiguous items results from an inherent dislike of obstruent-initial onsets (e.g., *taf* in *traf.**t**af*, relative to *raf* in *traf.**r**af*), then this dislike should be abolished for onsets that are comprised of obstruent-obstruent combinations (e.g., *ʃ vag.**v**ag* vs. *ʃ vag.**ʃ** ag*), as here, the congruent reduplicative affix also begins with an obstruent (e.g., *ʃ vag.**ʃ** ag*). But the acceptability of these items fully matched the omnibus pattern. In the relative rating, contiguous reduplicants (M=2.39) elicited reliably higher rating than both noncontiguous reduplicants (M=1.84; Z=4.26, p<.0001) and controls (M=1.78; Z=4.31, p<.0001), which, in turn did not differ (Z=0.41). Similarly, in absolute rating, contiguous reduplicants (M=3.03) elicited reliably higher rating than both noncontiguous reduplicants (M=2.32; Z=4.28, p<.0001) and controls (M=2.34; Z=4.63, p<.0001), which did not differ from each other (Z=0.48). These results make it clear that the preference for items like *trafraf* is inexplicable by the inherent properties of reduplicative forms — either structural or statistical. Having rejected these alternative explanations, we thus conclude that the preference for items like *trafraf* is specifically due to the effect of CONTIGUITY on the relation between the reduplicant and the base.

4. English speakers: Phonological forms (Experiments 2a-f)

The preference of Hebrew speakers for contiguous forms like *trafraf* is consistent with the hypothesis that morphological reduplication requires CONTIGUITY. However, the question arises whether CONTIGUITY targets the morphological level *selectively*. That is, whether speakers only favor contiguous forms when they assign the input a

morphological parse, but not when the same input is parsed as phonological identity.

To address this question, we next turn to English. English exhibits few forms of reduplication (see (3)), and speakers generalize them to new forms (Nevins & Vaux, 2003; Oden & Lopes, 1981; Parker, 2003; Pinker & Birdsong, 1979). Nonetheless, the experience of English speakers with reduplication is far more limited than that of Hebrew speakers. Furthermore, while Hebrew routinely uses reduplication in forming major lexical category (i.e. nouns, verbs, adjectives), in English, reduplication mostly concerns post-lexical and syntactic processes (Ghomeshi, Jackendoff, Rosen, & Russell, 2004), and consequently, it may not form part of the core morphology. Accordingly, it is unclear whether English speakers interpret novel lexical forms such as *trafraf* morphologically or phonologically — the existing findings do not address this question — and if they do interpret such forms morphologically, the question is whether CONTIGUITY is enforced.

(3) English reduplication[10]
   a. Dismissal reduplication (Nevins & Vaux, 2003):
      Metalinguistic-shmetalinguistic; reduplication-shmeduplication
   b. Full reduplication:
      bye-bye, pee-pee, no-no
   c. Rhyming:
      teenie-weenie, *itsy-bitsy*, *hoity-toity*
   d. Ablaut:
      chit-chat , ding-dong, zig-zag,
   e. Contrastive focus reduplication (Ghomeshi et al., 2004):
      Did you bring chicken salad or SALAD-salad

To address these questions, we next elicited acceptability ratings for a new set of word triplets, isomorphic to those used in the Hebrew experiments. We reasoned that, if the (limited) experience of English speakers with reduplication is sufficient to support a reduplicative morphological parse of the input, and if this parse is further constrained by CONTIGUITY, then the results of English speakers should mirror the Hebrew data. In contrast, if despite some experience with reduplicative forms, English speakers do not typically encode doubling morphologically, i.e., as reduplication, then by default, doubling will be parsed as purely phonological identity. Since phonology is not constrained by CONTIGUITY, and since phonological identity is generally dispreferred (by the OCP), we would then expect English speakers to consider *trafraf* no better than *traftaf*, and to disprefer them both compared to *trafkam*.

4.1. Methods

The materials consisted of 30 novel word triplets, exhibiting either contiguous reduplication (e.g., *trafraf*), noncontiguous reduplication (i.e., violations of the "no

[10] We thank two anonymous reviewers for these examples.

skipping" requirement, e.g., *traftaf*),[11] or no reduplication (e.g., *trafkam*). All items were phonotactically legal in English. To ensure a uniform syllabification of these items (as CCVC.CVC), we also designed the materials such that, in most items (75/90 items), the medial consonantal cluster formed an illegal English onset (see Appendix III), and consequently, the first consonant in the cluster was forced to the coda position (e.g., *snarnar* can only be syllabified as *snar.nar*; the syllabification **sna.rnar* is impossible due to the illicit *rn* onset). Experiments 2a-b elicited relative and absolute ratings (respectively). The procedure was identical to the Hebrew experiments, except that the materials were presented on the computer screen (rather than printed on paper, as in the Hebrew experiment), and participants rated their acceptability as potential English words. Two groups of native English speakers (N=30 each), students at Northeastern University, took part in these experiments.[12] With the exception of the specific rating response, the instructions to the two groups were identical. The instructions for the relative rating procedure are provided below.

"*In this experiment, you will be presented with triplets of printed words. The words do not exist in English, but some might sound better than others. We would like your opinion as to how they sound. In each set on the page, you will see three words in a column. Please sound out each word in your head, and then rank it relative to the other members of that set. Give rank 1 to the best word, 3 to the worst, and 2 to the intermediate one. Indicate your choice by writing the rank number on the line next to each word. Do not think too hard about it; just go with your gut reaction".*

4.2. Results and discussion

An inspection of the means (see Figure 2) suggests that English speakers were sensitive to the structure of the input, as they responded differentially to reduplicative items and the control. However, they showed no hint of CONTIGUITY effect.

[11] We chose these non-contiguous forms because we expected this violation of CONTIGUITY to be more salient than violations of the linearity requirement (e.g., *trafrat*). Although the results from Hebrew suggested that these two forms of non-contiguity produce similar outcomes, it remains to be seen whether this conclusion holds for English speakers.

[12] Second language information was obtained from 45 of the 60 participants in Experiment 2. Most participants were monolingual English speakers. There was a total of 10 participants who (in addition to native English competence) were also native speakers of another language: Gujarati (2), Korean (2), Arabic, Hindi, Japanese, Portuguese, Swedish, and Ukrainian.

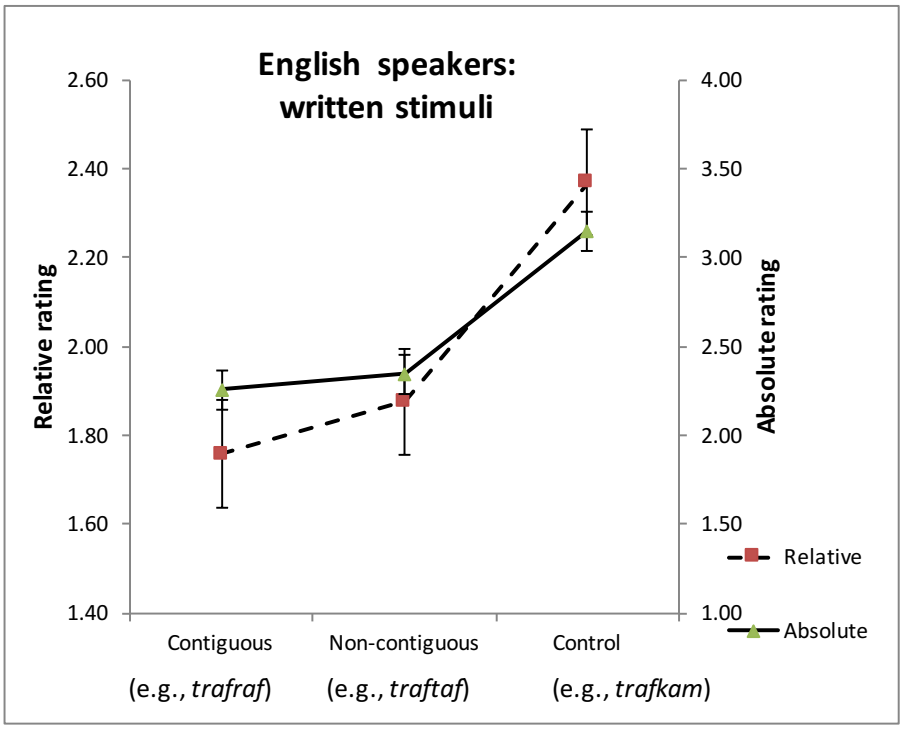

Figure 2. Acceptability ratings in Experiment 2a-b. Error bars are 95% confidence intervals for the difference between the means.

The indifference of English speakers to CONTIGUITY is unlikely due to their inability to encode the presence of doubling in surface forms (which is presumably necessary for the detection of reduplication). The Friedman nonparametric ANOVA yielded a significant effect of type in both the relative ($X^2_1(2)=17.96$, $p<.0002$; $X^2_2(2)=30.47$, $p<.0001$) and absolute ($X^2_1(2)=38.44$, $p<.0001$; $X^2_2(2)=45.79$, $p<.0001$) rating tasks. Moreover, Wilcoxon Matched Pairs Tests (see Table 3) showed that in each of the rating procedures, all reduplicative forms (*trafraf* and *traftaf*) elicited reliably *lower* ratings than non-reduplicative controls — a result that contrasts with the reduplication *preference* of Hebrew speakers. But while English speakers were sensitive to the presence of identical elements in the input, they were utterly indifferent to their CONTIGUITY.

Table 2. Pair-wise statistical contrasts in Experiment 2a-b.

				Participants		Items	
Exp.	N	**Procedure**	**Contrast**	***Z***	***p***	***Z***	***p***
2a	30	Relative rating (computer)	trafraf vs. traftaf	1.26	0.210000	2.31	0.030000
			trafraf vs. trafkam	3.19	0.002000	4.60	0.000010
			traftaf vs. trafkam	3.39	0.000700	4.10	0.000050
2b	30	Absolute rating (computer)	trafraf vs. traftaf	1.81	0.080000	1.44	0.160000
			trafraf vs. trafkam	4.55	0.000020	4.78	0.000001
			traftaf vs. trafkam	4.41	0.000020	4.78	0.000001

4.3. Replications and extensions

Before considering the theoretical significance of this finding, we first ensured that the insensitivity of English speakers to CONTIGUITY is a reliable observation that is inexplicable by methodological factors. Experiments 2c-f thus conducted several mini-replications of the original results with four new groups of participants. All participants were native English speakers, students at Northeastern University; the number of participants per experiment is provided in Table 3. The instructions are given in Appendix IV.

One possibility is that the divergence between English and Hebrew speakers is due to the use of different presentation modes (computer vs. printed pages, for English vs. Hebrew speakers respectively). To address this concern, we replicated the experiment in a paper and pencil version (Experiment 2c). The results (see Figure 3 and Table 3) were virtually unchanged (i.e., no CONTIGUITY preference).

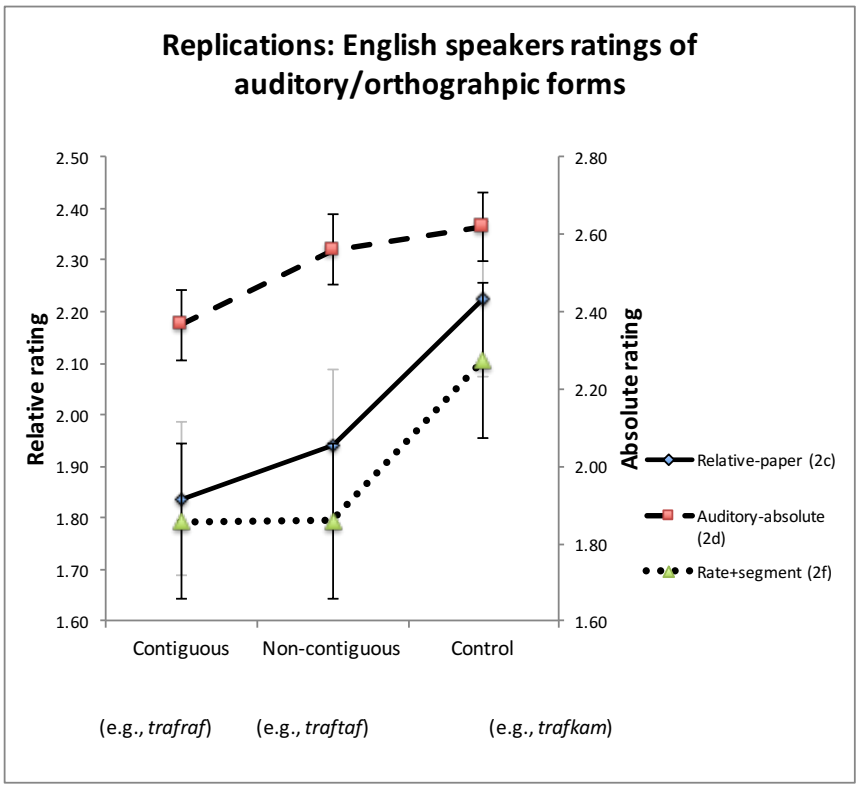

Figure 3. Acceptability ratings in Experiment 2c,d,f. Error bars are 95% confidence intervals for the difference between the means

Another possible source of English speakers' insensitivity to CONTIGUITY is their failure to decode the phonological form of our printed materials. We believe this concern is unlikely, given that skilled English readers are known to automatically extract phonological structure from print (e.g., Van Orden, Pennington, & Stone, 1990), but we nonetheless addressed this possibility. Experiment 2d thus elicited absolute ratings to an oral rendition of the same materials (uttered by a native

English speaker). No hint of a CONTIGUITY preference emerged (see Figure 3 and Table 3).

Table 3. Pair-wise statistical contrasts in the replications of Experiment 2(c-f)

				Participants		Items	
Exp.	N	**Procedure**	**Contrast**	***Z***	***p***	***Z***	***p***
2c	17	Relative rating (paper)	trafraf vs. traftaf	1.23	0.22	1.36	0.18
			trafraf vs. trafkam	1.96	0.05	3.74	0.0002
			traftaf vs. trafkam	1.59	0.12	2.47	0.02
2d	15	Absolute rating (auditory materials)	trafraf vs. traftaf	1.73	0.09	2.19	0.03
			trafraf vs.trafkam	2.27	0.03	3.07	0.003
			traftaf vs. trafkam	0.74	0.46	1.05	0.29
2f	17	Segment and rate (printed materials, absolute rating)	trafraf vs. traftaf	0.07	0.94	0.34	0.73
			trafraf vs. trafkam	1.70	0.09	3.96	0.002
			traftaf vs. trafkam	2.08	0.04	4.08	0.0002

We next attempted to direct participants' attention to the critical CONTIGUITY contrast by presenting them with pairs of matched reduplicative auditory stimuli — either contiguous or non-contiguous (e.g., *trafraf* vs. *traftaf*, with order counterbalanced) for a forced choice preference (Experiment 2e). Once again, English speakers (N=15) appeared entirely oblivious to CONTIGUITY: they selected the reduplicative items on 49% of the trials, at a rate that did not differ from chance ($t<1$ by participants and items).

Finally, as a last effort to elicit a CONTIGUITY preference, we attempted to call attention to the morphological base via syllabification. To this end, Experiment 2f presented yet another group of English speakers with the printed word-triplets for relative rating. Prior to rating, however, we asked the participants to first silently pronounce the printed string (to promote its phonological encoding), and then to parse each word onto syllables and indicate their response on a printed page (by marking the location of the syllable boundary with a line). An inspection of the segmentation responses suggested that on most trials (M=93%), the right edge of the base (CCVC) is aligned with the right edge of the first syllable (e.g., *traf⌋raf*); this response is as expected, given that we selected items where the middle consonant clusters cannot serve as a complex onset in English. Nonetheless, the rating of these forms yielded no evidence of CONTIGUITY. Ratings for *trafraf* and *traftaf* did not differ, and they were both dispreferred relative to *trafkam.*

5. English speakers: morphological forms (Experiment 3)

Our six experiments with English speakers make it clear that their indifference to the CONTIGUITY constraint is a reliable phenomenon. While Hebrew speakers enforce CONTIGUITY and favor reduplication, English speakers ignore CONTIGUITY and disfavor reduplicative to nonreduplicative forms.

These contrasts are readily explained by the hypothesis that English and Hebrew speakers assign different parses to these doubling forms. For Hebrew speakers, novel forms like *trafraf* are unambiguously morphological reduplication, just like 'orphan' reduplicative forms in the language, which likewise have no familiar base (e.g., *ʃravrav* (2c); Bat-El, 2006). As such, these forms must abide by CONTIGUITY, and they are thus better formed (relative to non-reduplicative forms). In the eyes of English speakers, however, *trafraf* only exhibits phonological identity (similar to *Trenton*). Because phonological identity is ill-formed due to OCP violation (see section 2), inputs like *trafraf* or *traftaf* are less acceptable than non-reduplicative controls. And since CONTIGUITY is inapplicable to phonological identity, English speakers are indifferent to the distinction between contiguous (*trafraf*) and non-contiguous (*traftaf*) forms. Of interest is whether English speakers indeed possess knowledge of CONTIGUITY, which they specifically apply to morphological reduplication.

To address this question, in our final, critical experiment, we once again presented the same materials to a (new) group of English speakers. As before, participants were asked to rate the three critical forms relative to each other, except that now, these items were implicitly presented as morphological diminutives — a morphological category marked by reduplication in many languages (Key, 1965). To this end, we first presented people with the base (e.g., *traf*) paired with a picture of a novel object. Participants were next presented with a miniature version of the same object, along with three printed options (e.g., *trafraf, traftaf, trafkam*), the precise same set of triplets from our previous English experiments (Experiment 2). They were asked to choose the best name for the miniature by rating the outcomes relative to each other. Previous research has shown that people are more likely to generalize a picture-word pairing to novel words that exhibit morphological affixation of the original base compared to ones that exhibit phonological changes (Bruening, Brooks, Alfieri, Kempe, & Dabasinskiene, 2012). Our question is whether speakers specifically require reduplicative morphological alternations to obey CONTIGUITY.

If English speakers lack knowledge of the CONTIGUITY constraint, then their indifference to contiguous forms should remain (i.e., *trafraf* vs. *traftaf*), irrespective of whether reduplication has phonological or morphological role. In contrast, if speakers possess knowledge of the CONTIGUITY constraint at the morphological level (between two morphologically related items), then once surface identity acquires a morphological interpretation, it will be parsed as reduplication, and consequently, the CONTIGUITY preference will now emerge.

5.1. Methods

A new group of 30 native English speakers took part in the experiment.[13] Materials consisted of the same set of 20 novel word triplets (Appendix III). Each such triplet

[13] Second language information was obtained from all the participants in Experiment 3. Most participants were monolingual English speakers. There was a total of 8 participants who (in addition

was paired with a single picture of a novel objects. In each trial, people saw a novel object, paired with a name (the base, e.g., *traf*), and they were asked to sound out its name. Next, they saw a diminutive object, along with the three word triplets (e.g., *trafraf, traftaf, trafkam*), in counterbalanced order. Participants were asked to sound out their names, and then choose the best name for the object. Prior to the main experimental session, participants were given three practice trials (with similar novel words), and invited to ask any questions. Participants received no feedback on their performance. The specific instructions are provided below.

"*In this experiment, you will be asked select the best name for objects pictured on the screen. In each trial, you will first learn a new word for an object. It is important that you sound out each word in your head and remember it before pressing space bar to continue. Then you will see a new object that is related to the first, and three possible names for it. Again, sound out each name in your head. Then pick which word is the best name for that object. Please select which name is the best, which is in the middle, and which is worst. To indicate your response, select the appropriate number beneath each word. You can change your answer by clicking on a yellow box to deselect it. There will be a practice section before the main experiment starts to help you understand the task. Do you have any questions*?"

5.2. Results

An inspection of the results (see Figure 4) shows that the simple change in procedure had a dramatic effect on the rating preferences. First, participants now exhibited a highly reliable preference for contiguous reduplicants relative to either the non-contiguous reduplicants or the non-reduplicative condition. In addition, the previous dislike of reduplication was now changed to an overall preference for the reduplicative forms (contiguous and noncontiguous) relative to the non-reduplicative controls.

These conclusions are borne out by the significant main effect of word type in the Friedman nonparametric ANOVA ($X^2_1(2)=16.79$, $p<.0003$; $X^2_2(2)=60.00$, $p<.0001$). The CONTIGUITY preference was significant relative to both the noncontiguous ($Z=3.33$, $p<.001$; $Z=4.78$, $p<.0001$) and control conditions ($Z=3.45$, $p<.0001$; $Z=4.78$, $p<.0002$). In addition, non-contiguous items (e.g., *traftaf*) were now more acceptable than controls (e.g., *trafkam*, ($Z=2.58$, $p<.02$; $Z=4.78$, $p<.0001$).

---

to native English competence) were native speakers of another language: Chinese (4), Korean (2), Arabic (1) Portuguese (1).
Most participants were monolingual; four participants reported speaking Chinese at home, two spoke Korean, one spoke Arabic and one Portuguese.

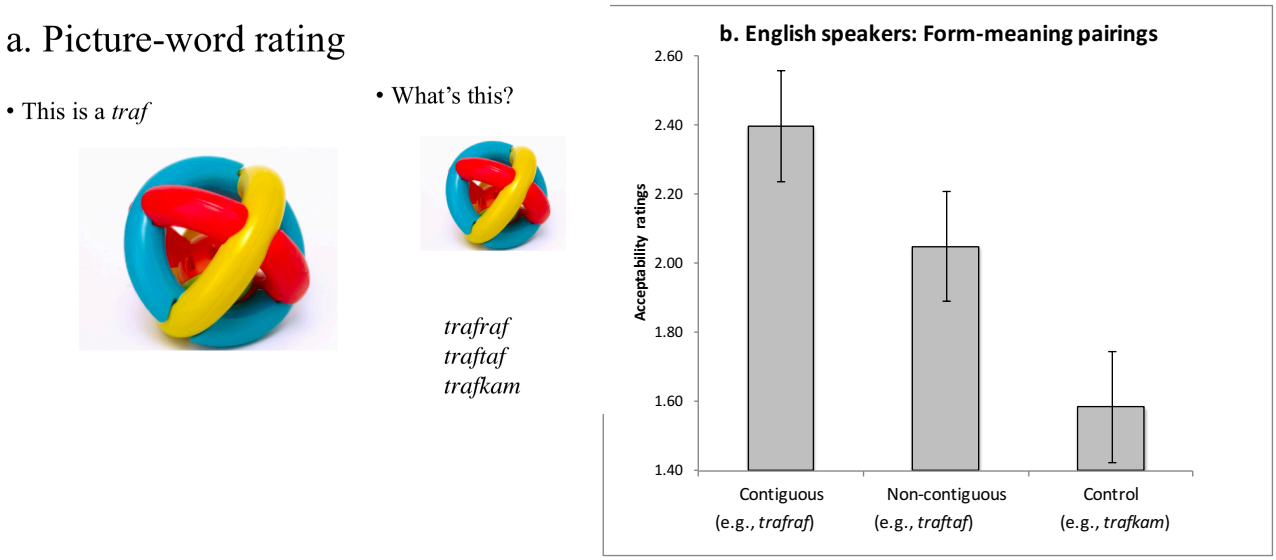

Figure 4. The picture-word rating procedure (a); and acceptability ratings in Experiment 3 (b). Error bars are 95% confidence intervals for the difference between the means.

It is unlikely that the effect of the form-picture pairing occurred at a purely phonological level (e.g., by calling attention to *traf* as a phonological constituent of *trafraf*) — recall that the syllabification task, described in Experiment 2f, yielded no hint of a CONTIGUITY preference despite the fact that, when asked to segment the word, people selected the base (e.g., *traf*) as the first syllable. We thus conclude that the semantic relation (diminutive) between the base and the probe words promoted their interpretation as morphologically complex. Once doubling was assigned a morphological role, it was parsed as reduplication, hence, it no longer violated the constraint on phonological identity (OCP).[14] And since reduplication reduces the number of phonological elements (by parsing doubling as copies), reduplicative forms (either contiguous or non-contiguous) are in fact better formed than non-reduplicative ones, hence, preferred.

We should note that the preference for the reduplicative forms (relative to controls) was obtained despite the violation of CONTIGUITY (e.g., for *traftaf*) — a result that differs from the Hebrew findings, where *traftaf* and *trafkam* were rated alike. This difference could have occurred either because of the different task demands, or because Hebrew speakers impose a higher penalty on CONTIGUITY violation (perhaps due to the prevalence of contiguous reduplicated forms in their language); further research is required to evaluate these possibilities. Crucially, once speakers assigned the input a morphological parse, the effect of the CONTIGUITY constraint spontaneously emerged in both languages.

[14] Our analysis assumes that the OCP constrains identical elements within the same morpheme (see footnote 1). Since reduplicative elements span different morphemes, their phonological form does not violate the OCP. Further studies should address the role of OCP in reduplicative forms.

6. General Discussion

Duality of patterning is a universal design feature of human language. Here, we asked whether the two levels of patterning — that of meaningless and meaningful elements — are each subject to distinct sets of constraints. To address this question, we exploited the structural ambiguity of doubling. Across languages, doubling is subject to two conflicting structural parses, as either phonological identity (as in the English *banana*) or morphological reduplication (as in the Manam *panana* 'chase', from the base *pana* 'run'), and these parses are each subject to distinct constraints. Morphological reduplication is subject to CONTIGUITY, whereas phonological identity is invariably dispreferred due to the OCP. Our experiments examined whether speakers doubling preferences shift depending on the level of analysis — phonology or morphology.

Results showed that, by default, English speakers parse bare forms like *trafraf* (and *traftaf*) as purely phonological representation, where identical elements are invariably dispreferred, irrespective of CONTIGUITY. But once the same surface forms are presented with reference to a morphological base (e.g., *trafraf* is the diminutive form of *traf*), English speakers favor a representation of reduplication, and spontaneously enforce CONTIGUITY (preferring *trafraf* over *traftaf*), as did speakers of Hebrew — a language in which morphological reduplication is pervasive.

These findings demonstrate that the constraints on doubling are selective with respect to the linguistic level of patterning — phonology vs. morphology. Furthermore, the parse projected by participants to doubled elements is constrained by speakers' linguistic experience. Given abundant experience with morphological reduplication, Hebrew speakers tend to interpret novel strings morphologically, even in the absence of any specific form-meaning associations. In contrast, English speakers, whose experience with reduplication is far reduced, tend to view the same strings as purely phonological forms, and they require explicit demonstration of form-meaning links in order to entertain a reduplicative parse. Crucially, once the morphological context is established, the effect of the CONTIGUITY constraint immediately emerges.

What is the source of these constraints? Optimality theory (Prince & Smolensky, 1993/2004) and the theory of Prosodic Morphology (McCarthy & Prince, 1998) assert that the OCP and CONTIGUITY are universal grammatical constraints. In line with this possibility, a growing body of experimental evidence suggests that people might exhibit knowledge of putatively universal grammatical constraints that are unattested in their language (e.g., Berent, Steriade, Lennertz, & Vaknin, 2007; Culbertson, Smolensky, & Legendre, 2012; Culbertson & Adger, 2014; Finley & Badecker, 2009; Gibson et al., 2013; Moreton, 2002). Our present results, however, do not allow us to evaluate the universality of the OCP and CONTIGUITY. Since each of our two participant groups has had at least some experience with reduplication, it is conceivable that participants induced these constraints from experience with their native language. Hebrew clearly presents speakers with ample evidence for CONTIGUITY. The English situation, however, is less certain. Unlike Hebrew, English reduplication does not form part of the core lexicon (Ito & Mester, 1999), so its relevance to lexical level reduplication (e.g., *trafraf*) is unclear.

Although our results cannot determine with certainty whether the OCP and CONTIGUITY are active universally, in the grammar of every speaker, there is nonetheless much linguistic evidence to suggest that these constraints apply in many languages (e.g., Berent & Shimron, 1997; Berent et al., 2001; Berent, Marcus, Shimron, & Gafos, 2002; Berent, Vaknin, & Marcus, 2007; Berkley, 1994; Boll-Avetisyan & Kager, 2014; Buckley, 1997; Domahs, Kehrein, Knaus, Wiese, & Schlesewsky, 2009; Frampton, 2009; Frisch et al., 2004; Frisch & Zawaydeh, 2001; Inkelas & Zoll, 2005; Kager, 1999; Kawahara et al., 2006; Marantz, 1982; McCarthy, 1979; McCarthy & Prince, 1995b; Raimy, 2012; Suzuki, 1998; Walter, 2007; Wilbur, 1973; Zuraw, 2002). The question then arises: why do languages converge on these particular restrictions?

Functional pressures could certainly provide part of the explanation. And indeed, doubling can present both costs and benefits. On the one hand, doubling is a liability, because it imposes known processing challenges ranging from repetition blindness in perception (Kanwisher, 1987, see also Frisch et al., 2004; Pierrehumbert, 1993), to biomechanical constraints on speech production (Walter, 2007), and lexical competition (Cohen-Goldberg, 2012). So all things being equal, functional pressures might render doubling dispreferred by the phonology (e.g., by the OCP). At the level of morphology, however, doubling can present a relative advantage. First, doubling underscores the form-meaning links between the base (e.g., *traf*) and the reduplicative form (e.g., *trafraf*), especially if the integrity of the base is preserved, as required by CONTIGUITY. Furthermore, the cost of doubling is generally weaker across morphemes (e.g., between a base and suffix, as in *raided*; Cohen-Goldberg, 2013; Cohen-Goldberg, Cholin, Miozzo, & Rapp, 2013). So while any form of doubling incurs sensory/motor costs, at the level of the morphology, these costs are offset by its advantages in marking form-meaning links. CONTIGUITY helps maximize those morphological gains.

The above scenario suggests that CONTIGUITY and the OCP are each functionally adaptive at the specific level of analysis at which they each operate. The OCP is a sensible phonological constraint because it mitigates against the known costs of repetition at the sensorimotor levels. By contrast, at the level of morphology, doubling costs are lower, and they are outweighed by the benefits incurred by CONTIGUITY, as this constraint underscores the links between form and meaning. This analysis converges with a large literature that explores the correspondence between the organization of the language system and functional pressures (e.g., Archangeli & Pulleyblank, 1994; Gibson et al., 2013; Hayes, Kirchner, & Steriade, 2004; Moreton & Pater, 2012; Pierrehumbert, 1993; Stampe, 1973; Steriade, 2001). The conclusion emerging from these studies is that many linguistic constraints are functionally grounded. But precisely how these functional pressures shape linguistic preferences is open to multiple interpretations (see (4)).

(4) The role of functional pressures in shaping linguistic preferences.
   a. *Direct functional account*:
      Functional pressures → linguistic behavior
   b. *Grammatical grounding:*
      Functional pressures → Universal grammar → linguistic behavior

One possibility is that functional constraints mold linguistic preferences *directly* (see (4a)). Contiguous forms (e.g., *trafraf*), in this view, are preferred not because they abide by linguistic constraints. Rather, *trafraf* is preferred because its perception and production is easier than *traftaf*. Put simply, the constraints on doubling are not linguistic — they are sensorimotor.

This account, however, faces two major challenges. First, our results show that a single linguistic form triggers *opposite* responses (aversion vs. preference) depending on the level of linguistic analysis (phonology vs. morphology). Subsequent research from our lab has also documented the converse pattern (Berent, Bat-El, Brentari, Dupuis, & Vaknin-Nusbaum, 2016). We found that English speakers spontaneously extend the *same* doubling preferences to novel signs in American Sign Language: signs with phonological doubling are systematically disliked, but once doubling is presented as a morphological operation (e.g., plurality), the doubling aversion shifts into a reliable preference. Together, these results outline a double dissociation between the sensorimotor demands of the stimulus and its acceptability: a single stimulus elicits opposite reactions, yet people's reactions are invariant to radical changes to stimulus modality (speech vs. sign). It is unclear how functional pressures would account for these findings.

A second (related) challenge to the direct functional account is presented by the distinct computational characteristics of the phonological and sensory/motor levels. The direct functional account requires that the representations at the linguistic and sensory/motor levels are commensurable — they share the same representational format, and abide by similar type of combinatorial principles. But there is reason to believe that linguistic and sensorimotor principles are computationally incommensurable. Most linguistic theories view the grammar as a discrete, algebraic system (Chomsky & Halle, 1968; Pinker & Prince, 1988; Smolensky & Legendre, 2006), whereas the analog and continuous representations that inform sensation and motor action undergo blending (Abler, 1989). So while there is much evidence to suggest that CONTIGUITY and OCP are functionally motivated, it is doubtful that these constraints can be reduced to functional pressures.

On an alternative formulation of the functional approach (see (4)b), functional pressures shape the language system, but these effects occur not directly, in on-line perception and action, but rather off-line and indirectly — in ontogeny and phylogeny. In this second view, the human language system satisfices functional pressures by favoring grammatical systems that are functionally grounded. The constraints operating within the grammar are by hypothesis, algebraic, and distinct from the sensory and motor system, but these grammatical constraints "conspire" to favor the computation of structures that are functionally adaptive. Thus, the grammar optimizes functional pressures using algebraic means (Berent, 2013).

Summarizing, doubling preference could originate from either grammatical or non-grammatical constraints, and these constraints could be either induced from experience or innately specified. While the source of doubling restrictions remains unknown, our present findings demonstrate that phonology and morphology are each subject to distinct sets of constraints. As such, these conclusions suggest the

morphology is an autonomous component of the language system, distinct from the phonology. Why duality of patterning has evolved, and how it is grounded in functional pressures are crucial questions for future research.

References

Abler, W. L. (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Systems, 12*, 1-13.

Archangeli, D., & Pulleyblank, D. (1994). *Grounded Phonology*. Cambridge MA: MIT Press.

Becker, M., & Bolozky, S. (2006). *Living lexicon of Hebrew Nouns*. Manuscript submitted for publication.

Berent, I., & Shimron, J. (1997). The representation of Hebrew words: Evidence from the Obligatory Contour Principle. *Cognition, 64*, 39-72.

Berent, I., Everett, D. L., & Shimron, J. (2001). Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology, 42*, 1-60.

Berent, I., Marcus, G. F., Shimron, J., & Gafos, A. I. (2002). The scope of linguistic generalizations: evidence from Hebrew word formation. *Cognition, 83*, 113-139.

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition, 104*, 591-630.

Berent, I., Vaknin, V., & Marcus, G. (2007). Roots, stems, and the universality of lexical representations: Evidence from Hebrew. *Cognition, 104*, 254-286.

Berent, I. (2013). *The phonological mind*. Cambridge: Cambridge University Press.

Berent, I., Bat-El, O., Brentari, D., Dupuis, A., & Vaknin-Nusbaum, V. (2016). The double identity of linguistic doubling. *Proceedings of the National Academy of Sciences, 113*, 13702-13707.

Berkley, D., Milam. (2000). *Gradient obligatory contour principle effects.* Northwestern University.

Berkley, D. M. (1994). The OCP and gradient data. *Studies in the Linguistic Sciences, 24*, 59-72.

Boll-Avetisyan, N., & Kager, R. (2014). OCP-PLACE in Speech Segmentation. *Language and Speech, 57*, 394-421.

Bolozky, S. (1999). *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: Brill.

Bruening, P. R., Brooks, P. J., Alfieri, L., Kempe, V., & Dabasinskiene, I. (2012). Children's Tolerance of Word-Form Variation. *Child Development Research, 2012*, 12.

Buckley, E. (1997). Tigrinya root consonants and the OCP. *U. Penn working papers in linguistics, 4*, 19-50.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 282-333). Cambridge: Cambridge University Press.

Cohen-Goldberg, A. (2013). Towards a theory of multimorphemic word production: The heterogeneity of processing hypothesis. *Language and Cognitive Processes, 28*, 1036-1064.

Cohen-Goldberg, A. M. (2012). Phonological competition within the word: Evidence from the phoneme similarity effect in spoken production. *Journal of Memory and Language, 67*, 184-198.

Cohen-Goldberg, A. M., Cholin, J., Miozzo, M., & Rapp, B. (2013). The interface between morphology and phonology: exploring a morpho-phonological deficit in spoken production. *Cognition, 127*, 270-286.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition, 122*, 306-329.

Culbertson, J., & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences, 111*, 5842-5847.

Domahs, U., Kehrein, W., Knaus, J., Wiese, R., & Schlesewsky, M. (2009). Event-related potentials reflecting the processing of phonological constraint violations. *Language And Speech, 52*, 415-435.

Everett, D. L., & Berent, I. (1999). An Experimental Approach to the OCP: Evidence for Violable Identity Constraints in Hebrew Roots. *Rutgers Optimality Archive*, Available http://ruccs.rutgers.edu/roa.html (ROA-235).

Feldman, L. B., & Bentin, S. (1994). Morphological analysis of disrupted morphemes: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology, 47*, 407-435.

Finley, S., & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language, 61*, 423-437.

Frampton, J. (2009). *Distributed Reduplication*. Cambridge, MA: MIT Press.

Frisch, S., A., Pierrehumbert, J., B., & Broe, M. B. (2004). Similarity avoidance and the OCP. *Natural Language and Linguistic Theory, 22*, 197-228.

Frisch, S. A., & Zawaydeh, B. A. (2001). The psychological reality of OCP-place in Arabic. *Language, 77*, 91-106.

Frost, R., Deutsch, A., & Forster, K. I. (2000). Decomposing morphologically complex words in a nonlinear morphology. *J Exp Psychol Learn Mem Cogn, 26*, 751-765.

Ghomeshi, J., Jackendoff, R., Rosen, N., & Russell, K. (2004). Contrastive Focus Reduplication in English (The Salad-Salad Paper). *Natural Language & Linguistic Theory, 22*, 307-357.

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A Noisy-Channel Account of Crosslinguistic Word-Order Variation. *Psychological Science, 24*, 1079-1088.

Hayes, B., Kirchner, R. M., & Steriade, D. (2004). *Phonetically based phonology*. Cambridge ; New York: Cambridge University Press.

Hockett, C. F. (1960). The origin of speech. *Sci Am, 203*, 89-96.

Inkelas, S., & Zoll, C. (2005). *Reduplication: Doubling in Morphology*. Cambridge, UK: Cambridge University Press.

Inkelas, S. (2008). The dual theory of reduplication *Linguistics, 46*, 351-401.

Ito, J., & Mester, A. (1999). The structure of the phonological lexicon. In N. Tsujimura (Ed.), *A Handbook of Japanese Linguistics* (pp. 62-100). Oxford: Blackwell.

Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Kanwisher, N. G. (1987). Repetition blindness: type recognition without token individuation. *Cognition, 27*, 117-143.

Kawahara, S., Ono, H., & Sudo, K. (2006). Consonant co-occurrence restrictions in Yamato Japanese. In T. Vance & K. Jones (Eds.), *Japanese/Korean Linguistics* (Vol. 14, pp. 27-38.). Stanford: CSLI Publications.

Key, H. (1965). Some Semantic Functions of Reduplication in Various Languages. *Anthropological Linguistics, 7*, 88-102.

Klein, S. (2015). *Frequency or Theory? The acquisition of consonants in Modern Hebrew*. Manuscript submitted for publication.

Leben, W. (1973). *Suprasegmental phonology*. Cambridge, MA: MIT press.

Lichtenberk, F. (1983). A Grammar of Manam. *Oceanic Linguistics Special Publications*, i-647.

Marantz, A. (1982). Re Reduplication. *Linguistic Inquiry, 13*, 435-482.

Marslen-Wilson, S. B. W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language & Cognitive Processes, 20(1) , p. 207-260*.

McCarthy, J. J. (1979). *Formal problems in Semitic phonology and morphology. Doctoral dissertation, MIT.* New York, 1985: Garland Press.

McCarthy, J. J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry, 12*, 373-418.

McCarthy, J. J., & Prince, A. (1995a). Faithfulness and reduplicative identity. In J. Beckman, S. Urbanczyk & L. Walsh Dickey (Eds.), *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory* (pp. 249-384).

McCarthy, J. J., & Prince, A. (1995b). Prosodic morphology. In J. A. Goldsmith (Ed.), *Phonological theory* (pp. 318-366). Oxford: Basil Blackwell.

McCarthy, J. J., & Prince, A. (1998). Prosodic morphology. In A. Spencer & A. Zeicky, M (Eds.), *Handbook of Morphology* (pp. 283-305). Oxford: Basil Blackwell.

Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition, 84*, 55-71.

Moreton, E., & Pater, J. (2012). Structure and substance in artificial‚Äêphonology learning, part I: Structure. *Language and Linguistics Compass, 6*, 686-701.

Nevins, A., & Vaux, B. (2003). Metalinguistic, shmetalinguistic: The phonology and morphology of shm-reduplication. *Proceedings from the Annual Meeting of the Chicago Linguistic Society, 39*, 702-721.

Oden, G. C., & Lopes, L. L. (1981). Preference for Order in Freezes. *Linguistic Inquiry, 12*, 673-679.

Ohala, D. K. (1999). The influence of sonority on children's cluster reductions. *Journal Of Communication Disorders, 32*, 397-421.

Parker, S. (2003). The psychological reality of sonority in English. *WORD, 54*, 359-399.

Pierrehumbert, J. (1993). *Dissimilarity in Arabic verbal roots.* Paper presented at the Proceedings of NELS 23, GLSA, Departments of Linguistics, University of Massachusetts, Amherst, MA.

Pinker, S., & Birdsong, D. (1979). Speakers' sensitivity to rules of frozen word order. *Journal of Verbal Learning and Verbal Behavior, 18*, 497-508.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*, 73-193.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell Publishing.

Raimy, E. (2012). *The Phonology and Morphology of Reduplication*.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review, 11*, 1090-1098.

Rose, S., & Walker, R. (2004). A typology of consonant agreement as correspondence. *Language, 80*, 475-531.

Smolensky, P., & Legendre, G. (2006). Principles of integrated connectionist/symbolic cognitive architecture. In P. Smolensky & G. Legendre (Eds.), *The harmonic mind: From neural computation to Optimality-theoretic grammar* (Vol. 1: Cognitive architecture, pp. 63-97). Cambridge, MA: MIT Press.

Stampe, D. (1973). *A dissertation on natural phonology.* The University of Chicago, Chicago.

Stemberger, J., & Treiman, R. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language, 25*, 163-180.

Steriade, D. (2001). *The phonology of perceptibility effects: The P-map and its consequences for constraint organization*. Manuscript submitted for publication.

Suzuki, K. (1998). *A typological investigation of dissimilation.* University of Arizona, Tucson, AZ.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning & Verbal Behavior, 14*, 638-647.

Ussishkin, A. (1999). The inadequacy of the consonantal root: Modern Hebrew denominal verbs and Output-Output correspondence. *Phonology, 16*, 441-442.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97*, 488-522.

Walter, M. A. (2007). *Repetition avoidance in human language.* Massachusetts Institute of Technology, Cambridge.

Wilbur, R. B. (1973). *The phonology of reduplication.* Unpublished Ph.D., University of Illinois at Urbana-Champaign, Ann Arbor.

Zuraw, K. (2002). Aggressive reduplication. *Phonology, 19*, 395-439.

Appendix I: The formal definition of CONTIGUITY

Formally, CONTIGUITY is defined by two constraints on the correspondence between the base and the reduplicant (McCarthy & Prince, 1995). One constraint (I-CONTIGUITY) requires that the material in the reduplicant (R) is copied without skipping (deletion); another constraint (O-CONTIGUITY) bans the addition (epenthesis) of new material.

a. I-CONTIG (“No Skipping”)
   The portion of S_1 standing in correspondence forms a contiguous string.
   Domain (**R**) is a single contiguous string in S_1.

b. O-CONTIG (“No Intrusion”)
   The portion of S_2 standing in correspondence forms a contiguous string.
   Range(**R**) is a single contiguous string in S_2.

A related constraint, LINEARITY, further requires that the order of the segments in the base be maintained, so *trafraf* (which obeys LINEARITY) is preferred to *trafrat* (which violates LINEARITY). In what follows, we assume that contiguous reduplicants must satisfy both constraints (CONTIGUITY and LINEARITY)

LINEARITY—“No Metathesis”
   S_1 is consistent with the precedence structure of S_2, and vice versa.
   Let $x, y \in S_1$ and $x', y' \in S_2$.
   If x **R** x' and y **R** y', then
   $x < y$ iff $\neg (y' < x')$.

Appendix II: The materials used in the Hebrew reduplication experiment[15]

Contiguous reduplication		Non-Contiguous reduplication		No reduplication	
שבגבג	ʃ vagvag	שבגשג	ʃ vagʃ ag	שבגרם	ʃ vagram
גבדבד	gvadvad	גבדגד	gvadgad	גבדמש	gvadmaʃ
דבנבן	dvanvan	דבנדב	dvandav	דבנפק	dvanfak
זגבגב	zvagvag	זגבזג	zgavzag	זגבקם	zgavkam
שגדגד	ʃ gadgad	שגדשג	ʃ gadʃ ag	שגדמף	ʃ gadmaf
קדבדב	kdavdav	קדבקב	kdavkav	קדברג	kdavrag
פדנדן	pdandan	פדנפד	pdanfad	פדנגס	pdangas
גזבזב	gzavzav	גזבגז	gzavgaz	גזברק	gzavrak
פזקזק	pzakzak	פזקפק	pzakfak	פזקלב	pzaklav
שחגחג	ʃ xagxag	שחגשג	ʃ xagʃ ag	שחגנל	ʃ xagnal
קחדחד	kxadxad	קחדקד	kxadkad	קחדלף	kxadlaf
קלמלם	klamlam	קלמקם	klamkam	קלמרש	klamraʃ
גלפלף	glaflaf	גלפגף	glafgaf	גלפנץ	glafnac
קלצלץ	klaclac	קלצקץ	klackac	קלצבר	klacvar
סמגמג	smagmag	סמגסג	smagsag	סמגרב	smagrav
בנסנס	bnasnas	בנסבס	bnasbas	בנסלם	bnaslam
פסמסם	psamsam	פסמפם	psamfam	פסמגן	psamgan
פעגעג	p?ag?ag	פעגפג	p?agfag	פעגרק	p?agrak
שפגפג	ʃ fagfag	שפגשג	ʃ fagʃ ag	שפגנב	ʃ fagnav
פקנקן	pkankan	פקנפק	pkanfak	פקנלש	pkanlaʃ
שפדפד	ʃ fadfad	שפדפש	ʃ fadfaʃ	שפדרך	ʃ fadrax
קפנפן	kfanfan	קפנפק	kfanfak	קפנרס	kfanras
פצגצג	pcagcag	פצגצף	pcagcaf	פצגלב	pcaglav
פקזקז	pkazkaz	פקזקף	pkazkaf	פקזרג	pkazrag
קרדרד	kradrad	קרדרק	Kradrak	קרדלף	kradlaf
גרטרט	gratrat	גרטרג	gratrag	גרטמש	gratmaʃ
דרצרץ	dracrac	דרצרד	dracrad	דרצרף	dracraf
גשתשת	gʃ atʃ at	גשתשג	gʃ atʃ ag	גשתלב	gʃ atlav
פתגתג	ptagtag	פתגתף	ptagtaf	פתגרד	ptagrad
קבתבת	kvatvat	קבתבק	kvatvak	קבתלש	kvatlaʃ
סגטגט	sgatgat	סגטגס	sgatgas	סגטמף	sgatmaf
בגשגש	bgaʃ gaʃ	בגשגב	bgaʃ gav	בגשנד	bgaʃ nad
פזכזך	pzaxzax	פזכזף	pzaxzaf	פזכרם	pzaxram
סחמחם	sxamxam	סחמחס	sxamxas	סחמרג	sxamrag
פטבטב	ptavtav	פטבטף	ptavtaf	פטבלק	ptavlak
בסנסן	bsansan	בסנסב	bsansav	בסנרץ	bsanrac
צפמפם	cfamfam	צפמפץ	cfamfac	צפמרב	cfamrav
סקמקם	skamkam	סקמקס	skamkas	סקמרב	skamrav
פקצקץ	pkackac	פקצקף	pkackaf	פקצרת	pkacrat
בשמשם	bʃ amʃ am	בשמשב	bʃ amʃ av	בשמגד	bʃ amgad

[15] (i) We use /c/ to represent the coda of ca***ts***, /ʃ / to indicate the onset of ***sh****ip.*
(ii) Hebrew *p, b* and *k* are often spirantizaed to *p*, *v*, and *x* respectively in post-vocalic position, where the difference in the script is in the presence vs. absence of a dot within the letter (e.g., בּ for *b* vs. ב for *v*). SInce spirantization displays a great degree of variation and that reading material is usually without diacritics, we did not add the diacritic in the experiment. Therefore, גבדבד, for example, could be read as either *gvadvad* or as *gvadbad*. In the above list we present the more common form, the one preserving identity (and thus violating spirantization).
(iii) Note that *c*, *k*, *p*, *m*, and *n* are represented by different letters in word final vs. nonfinal positions. In the reduplicated form *klaclac* קלצלץ, for example, the segment *c* and its copy get different letters — ץ in final position and צ eleswhere.

Appendix III: The materials used in the English reduplication experiment

Contiguous reduplication	Non-Contiguous reduplication	No reduplication
blaflaf	blafbaf	blaftak
blavlav	blavbav	blavmar
brafraf	brafbaf	brafgat
bravrav	bravbav	bravgat
drafraf	drafdaf	drafpag
drakrak	drakdak	drakmav
dravrav	dravdav	dravkam
flaslas	flasfas	flaspar
fralral	fralfal	fralgad
frasras	frasfas	frasmal
glanlan	glangan	glanvap
glatlat	glatgat	glatrab
glavlav	glavgav	glavdap
gravrav	gravgav	gravlat
klaflaf	klafkaf	klafpar
kravrav	kravkav	kravlan
plaflaf	plafpaf	plafsav
prafraf	prafpaf	praftak
slaflaf	slafsaf	slafmak
slanlan	slansan	slanvak
slavlav	slavsav	slavnag
smafmaf	smafsaf	smafkal
smalmal	smalsal	smalgar
smarmar	smarsar	smarvak
smavmav	smavsav	smavgar
snafnaf	snafsaf	snafgab
snarnar	snarsar	snarkal
snavnav	snavsav	snavmak
trafraf	traftaf	trafkam
travrav	travtav	travgam

Appendix IV: Instructions for Experiments 2c-f

Experiment 2c: relative rating (paper and pencil)

In this experiment, you will be presented with triplets of printed words. The words do not exist in English, but some might sound better than others. We would like your opinion as to how they sound.

In each set on the page, you will see three words in a column. Please sound out each word in your head, and then rank it relative to the other members of that set. Give rank 1 to the best word, 3 to the worst, and 2 to the intermediate one. Indicate your choice by writing the rank number on the line next to each word. Do not think too hard about it; just go with your gut reaction.

After you are finished ranking all three words in a set, go on to the next set to the right, and then on to the next row and so on. Do you have any questions?

Experiment 2d: Absolute rating (auditory words)

In this experiment, you will hear one word at a time. These words do not exist in English, but some might sound better than others. Your job is to rate them, 1 to 5, with five being the extremely good and one being very bad.

You will press space bar to start each trial, and then the word will play. Please be sure to listen carefully to each word. You can give your answer on the number pad. Please use the whole range of the scale as you judge each word. Do not overthink the answer; just go with your gut reaction.

There will be a practice section before the main experiment starts to help you understand the task. Do you have any questions?

Experiment 2e: rate pairs of auditory words

In this experiment, you will hear pairs of words. The words do not exist in English, but some might sound better than others. We would like you to select which one sounds best in English.

You will press space bar to start each trial, and then the two words will play. To select the first one, press 1. To select the second, press two. Do not overthink the answer; just go with your gut reaction. Please be sure to listen carefully as the words are playing.

There will be a practice section before the main experiment starts to help you understand the task. Do you have any questions?

Experiment 2f: rate and segment printed words

In this experiment, you will be presented with triplets of printed words. The words do not exist in English, but some might sound better than others. We would like your opinion as to how they sound.

In each set on the page, you will see three words in a column. For each such triplet, we ask you to do three simple tasks. First, please sound out each word in your head. Then, please indicate how you might divide these words into part (e.g., *bloglog=blog+log, blog+bog, blo+glog*) by marking the boundary by a line. Finally, please rank it relative to the other members of that set. Give rank 1 to the best word, 3 to the worst, and 2 to the intermediate one. Indicate your choice by writing the rank number on the line next to each word. Do not think too hard about it; just go with your gut reaction.

After you are finished ranking all three words in a set, go on to the next set to the right, and then on to the next row and so on. Remember, first sound out the words, then mark the boundary, and finally rate them.
Do you have any questions?