Phonology **30** (2013) 253–295.
doi:10.1017/S0952675713000134

*Learning the identity effect as an artificial language : bias and generalisation**

Gillian Gallagher
New York University

The results of two artificial grammar experiments show that individuals learn a distinction between identical and non-identical consonant pairs better than an arbitrary distinction, and that they generalise the distinction to novel segmental pairs. These results have implications for inductive models of learning, because they necessitate an explicit representation of identity. While identity has previously been represented as root-node sharing in autosegmental representations (Goldsmith 1976, McCarthy 1986), or implicitly assumed to be a property that constraints can reference (MacEachern 1999, Coetzee & Pater 2008), the model of inductive learning proposed by Hayes & Wilson (2008) assumes strictly feature-based representations, and is unable to reference identity directly. This paper explores the predictions of the Hayes & Wilson model and compares it to a modification of the model where identity is represented (Colavin *et al*. 2010). The results of both experiments support a model incorporating direct reference to identity.

1 Introduction

Many languages distinguish between identical and non-identical segments with respect to some phonotactic pattern. For example, in Bolivian Aymara (Aymaran) pairs of identical ejectives may co-occur in a root, as in [t'ant'a] 'bread', but pairs of non-identical ejectives may not: *[t'ank'a]. This pattern is dubbed the 'identity effect' by MacEachern (1999). The classes of identical and non-identical segments have traditionally been analysed with suprasegmental, representational structures, either double linking in autosegmental phonology (Goldsmith 1976, McCarthy 1986, 1988) or reduplicative correspondence (Buckley 1990, Hudson 1995, Rose 1997, Gafos 1998). The sets of identical and

* I am particularly grateful to Peter Graff and three anonymous reviewers for detailed comments on this work, and also to Tal Linzen for helpful discussion. I also thank participants in the Fall 2012 phonology seminar and the Phonetics and Experiment Phonology reading group at New York University.

non-identical segments are thus represented differently from the featurally defined natural classes that are also commonly active in phonology, like the set of labial consonants (picked out by the feature [LABIAL]) or the set of stop consonants (picked out by the features [−continuant, −sonorant]). The identical/non-identical distinction falls out naturally given certain representational assumptions, but it is difficult to capture in models of inductive learning that assume strictly feature-based representations, as in the UCLA phonotactic learner (Hayes & Wilson 2008).

This paper explores the predictions of the UCLA phonotactic learner, an inductive model with no explicit representation of identity, and compares it to a modification of the model where identity is represented as copying (Colavin *et al.* 2010). In the model with no representation of identity, henceforth the BASELINE MODEL, the identity effect must be encoded as the cumulative effect of specific constraints on each non-identical pair of segments. Under this analysis, however, the identity effect is formally identical to any arbitrary distinction between segment combinations, and is not predicted to generalise to novel forms; cf. Marcus *et al.* (1999), who identify a similar failing in connectionist models, and argue that variables must be included in cognitive representations. A model in which identity is represented, henceforth the COPYING MODEL, allows for the identity effect to be learned as a general restriction on all pairs of non-identical segments, and thus predicts both that the identity effect is more general than other patterns and that it should extend to novel forms. Two artificial grammar experiments test the predictions of these two models, and demonstrate that participants' performance is better accounted for with the copying model. The results of Experiment 1 show that participants learn a pattern based on the distinction between identical and non-identical consonant pairs better than a pattern based on an arbitrary distinction between consonant pairs. Experiment 2 shows that participants generalise a distinction between identical and non-identical consonant pairs to novel pairs of identical and non-identical consonants. Both experiments support the analysis of the identity effect as a broad generalisation over identical and non-identical segments, evidencing the need for an explicit representation of identity in the phonological grammar.

The paper is organised as follows. §2 presents the typological pattern known as the identity effect in more detail and discusses different analyses of the pattern. In §3, the predictions of two inductive phonotactic models are compared, the baseline model proposed in Hayes & Wilson (2008) and the model proposed in Colavin *et al.* (2010), in which identity is represented as segmental copying. Experiment 1 is presented in §4, comparing learning of the identity effect to a formally identical, but arbitrary, pattern, and Experiment 2 is discussed in §5, testing whether the identical–non-identical distinction generalises to novel forms. §6 concludes.

2 The identity effect

2.1 The pattern

In many languages, identical segments are exempt from phonotactic restrictions that hold of non-identical segments. This is dubbed the 'identity effect' by MacEachern (1999). Consider the data from Bolivian Aymara in (1), taken from de Lucca's (1987) dictionary and discussed in MacEachern (1999) and Gallagher (2010a, b). Disyllabic roots in Bolivian Aymara may contain a single ejective (1a), but may not contain two ejectives (1b) unless they are identical (1c).

(1) a.	ʧ'aka	'bone'	b.	*ʧ'ak'a	c.	ʧ'aʧ'a	'to soak'
	k'apa	'cartilage'		*k'ap'a		k'aʎk'u	'acidic'
	p'eqe	'head'		*p'eq'e		p'ap'i	'type of fish'
	q'olti	'drink'		*q'olt'i		q'oq'e	'to catch fire'
	t'aku	'calm'		*t'ak'u		t'ant'a	'bread'

In other languages, like the varieties of Quechua spoken in Bolivia and the Southern Peruvian highlands, both non-identical and identical pairs of ejectives are absent from the language (e.g. *[k'…p'], *[k'…k']). Thus a phonotactic restriction on the co-occurrence of a class of sounds may apply uniformly to all segment pairs in a language, or it may target only non-identical pairs.

The identity effect is widely attested in languages with restrictions on laryngeally marked consonants. MacEachern finds that pairs of identical ejectives, aspirates or implosives are attested despite restrictions on non-identical pairs in Tz'utujil (Mayan), Hausa (Afro-Asiatic), Peruvian Aymara (Aymaran) and Gojri (Indo-Aryan). This same pattern is also found for ejectives in other Mayan languages, e.g. Chol (Gallagher & Coon 2009) and Yucatec Mayan (Straight 1976) and for implosives in Muna (Austronesian) (van den Berg & Sidu 1996, Coetzee & Pater 2008).

Restrictions on consonant co-occurrence may also apply equally to identical and non-identical segments, as mentioned above for ejectives in Quechua. This type of pattern is also found for ejectives in Shuswap (Salish) (MacEachern 1999) and Yapese (Austronesian) (Jensen 1977), and for pairs of aspirates in Quechua, Ofo (Siouan), Sanskrit (Indo-Aryan) and Souletin Basque (isolate) (MacEachern 1999).

Languages with restrictions on major place of articulation may also show the identity effect, whereby pairs of non-identical homorganic consonants may be underattested or unattested, while pairs of identical consonants are attested. The pattern is clearly illustrated in the co-occurrence of labial consonants in Muna: the attestation of a pair of labials generally decreases with the similarity of the two consonants, but identical consonants occur more often than expected by chance, despite being maximally similar (Coetzee & Pater 2008). The examples in (2) show pairs of non-identical and identical labial consonants and their observed/expected (O/E) ratios (taken from Coetzee & Pater 2008); a similar effect is found for coronals and dorsals. An O/E around 1 indicates that a pair of consonants occurs as often as expected by chance; below 1 indicates underattestation and above 1 overattestation (Pierrehumbert 1993, Frisch *et al*. 2004).

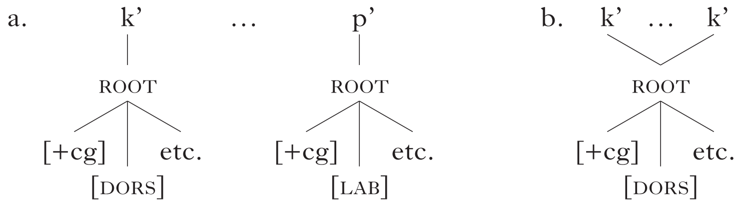
(2) *The identity effect in Muna: long-distance restriction on labials*

a.	m…f	1·04	b…p	0·10	b.	p…p	1·46
	b…f	0·58	p…f	0·07		b…b	2·79
	m…p	0·39	m…b	0·07		f…f	2·50
						m…m	1·24

A similar pattern is seen in the restrictions on homorganic consonants in Semitic (McCarthy 1986, Berent & Shimron 1997, Frisch & Zawaydeh 2001, Frisch *et al*. 2004), though in Semitic languages identical pairs of consonants are restricted to certain positions within the root, and are often analysed as deriving from a single consonant that is either linked to multiple prosodic positions (McCarthy 1986, 1988) or reduplicated (Gafos 1998). Ngbaka (Niger-Congo; Thomas 1963, Mester 1986) also shows this type of pattern, where homorganic, highly similar pairs of consonants are absent but identical pairs are widely attested, as do Javanese (Austronesian) and Bolivian Aymara (Graff & Jäger 2009).

2.2 Analysing the identity effect

Multiple proposals have been made for distinguishing between identical and non-identical segments in phonological patterns. The classic autosegmental account of the identity effect appeals to the idea of double linking. Under this line of analysis, identical segments are a single featural representation linked to two prosodic positions, and thus while on the surface there are two phonetically distinct segments, the phonology operates as if there is just one segment. This doubly linked representation allows identical segment pairs to escape violation of standardly formulated Obligatory Contour Principle (OCP) constraints (Leben 1973, Goldsmith 1976, McCarthy 1986, 1988, Coetzee 2009a). An OCP constraint takes the form $*[\alpha F][\alpha F]$, disallowing multiple instances of the same feature specification on the same autosegmental tier. For example, a restriction on two ejectives in a root results from a constraint on multiple instances of [+constricted glottis], the feature that distinguishes ejectives from other stops: *[+cg][+cg]. Pairs of identical segments can escape violation of an OCP constraint if a single root node is doubly linked to two consonantal positions, because there is only one [+constricted glottis] autosegment, as shown in (3). In (3a), there are two distinct root nodes and two instances of [+constricted glottis], which violates *[+cg][+cg]. In (3b), there is only a single root node and no sequence of [+constricted glottis] specifications, and thus no OCP violation.

(3)


The autosegmental account of the identity effect relies on a covert distinction between pairs of surface segments that are represented as distinct segments and those that are represented as a single segment. While the class of all identical segments is not easily definable in featural terms, double linking allows the grammar to reference identical segments as a class: identical segments are those that share a single root node.

Other analyses of the identity effect also assume that the identical or non-identical status of a pair of segments is something that constraints in Optimality Theory (Prince & Smolensky 1993) can reference. MacEachern (1999) analyses the identity effect as the interaction between an OCP constraint and a constraint requiring complete identity between consonants of a root, BeIdentical. Coetzee & Pater (2008) assume that OCP constraints generally apply only to non-identical segments; in languages where identical segments are disallowed, they are ruled out by a specific constraint against identical segments. Both of these proposals assume that the identity or non-identity of segments is something that constraints can refer to, just like any other feature value that may be shared or unshared between a pair of segments.

3 Constraint induction and the representation of identity

Any account of the identity effect as such relies on some representational distinction between identical and non-identical consonants, either in the phonological representation itself or in the structure of grammatical generalisations. In the baseline model of inductive phonotactic learning proposed in Hayes & Wilson (2008), which assumes a very basic structure for phonotactic generalisations, no such representational distinction is available. This section compares the predictions of the baseline model, with no explicit representation of identity, to the predictions of the copying model, where identity is represented (Colavin *et al.* 2010). The results of the two artificial grammar experiments in §4 and §5 support the predictions of a model that refers directly to identity, and thus provide an argument that an accurate model of learning must allow for an explicit distinction between identical and non-identical segments. The predictions of the baseline model are laid out schematically in §3.1, and contrasted with the copying model in §3.2. The models are compared concretely with learning simulations in §3.3.

3.1 The baseline model and predictions for identity

The principle behind the UCLA phonotactic learner (Hayes & Wilson 2008) as a model of phonotactics is to begin with a maximally simple theory of representations and constraints and see what range of phonological patterns can be learned with this baseline model. Evidence for enriching either representations or the structure of constraints comes from generalisations that cannot be stated or learned in the baseline model.

The UCLA phonotactic learner, the baseline model, starts with a feature set to distinguish the segments of a language and a formula for constructing constraints on feature combinations and sequences, and arrives at a grammar of weighted constraints, weighted via an algorithm based on the principle of Maximum Entropy (maxent; for background on maxent in general see Jaynes 1983, Manning & Schütze 1999 and Klein & Manning 2003, and for applications to phonological grammar Goldwater & Johnson 2003). The grammar assigns a probability distribution over the possible and impossible structures in a language, according to the formula in (4a). (4a) says that the probability of a form x, $P(x)$, is equivalent to the maxent value, defined as e raised to the negation of the score assigned by the grammar, $h(x)$, divided by Z, the sum of the total maxent values of all possible forms Ω. The score for a given form, $h(x)$, is defined in (4b); it is the sum of the product of the weights w of each constraint C that the form violates and the number of times the form violates that constraint. The total maxent values of all possible forms, Z, is defined in (4c).

(4) a. $P(x) = \frac{e^{-h(x)}}{Z}$ b. $h(x) = \sum_i w_i C_i(x)$ c. $Z = \sum_{x' \in \Omega} e^{-h(x')}$

Given a feature set that uniquely defines each segment, the learner constructs all possible natural classes, and then constructs all possible constraints from these natural classes. Constraints are negative statements that rule out natural classes, either in a single matrix penalising a type of segment (e.g. *[+cont, −son, LAB], penalising labial fricatives), or in a sequence of feature matrices, penalising sequences of segments (e.g. *[−cont, −son][−cont, −son], penalising adjacent sequences of stops).[1]

To arrive at a grammar, the learner selects a constraint from the set of possible constraints, and assigns the constraint a weight by testing it against the attested forms in the learning set. The goal of the weighting algorithm is to maximise the likelihood of attested forms and minimise the likelihood of unattested forms; thus a high weight is assigned to constraints that penalise unattested forms, and a low or zero weight is assigned to constraints that penalise some or only attested forms. The model

[1] Hayes & Wilson restrict their model to learning constraints with maximally three feature matrices, but this is a parameter in the model that can be stipulated.

first selects the most accurate constraints to add to the grammar, and then, within this set, selects the more general constraints. The accuracy of a constraint is defined roughly as the observed/expected ratio of violations on that constraint; it is the number of times the constraint is violated in the learning data divided by the number of violations that are expected given the current grammar (see Hayes & Wilson 2008: §4.2.1 for further discussion). The generality of a constraint is a function first of the number of feature matrices referred to in the constraint, and second of the number of features referred to in a given feature matrix. Shorter constraints are preferred over longer constraints (e.g. *[+voice, LAB] is preferred over *[+voice][+voice]), and among constraints of the same length, constraints that refer to fewer natural classes are preferred (e.g. *[+voice] is preferred over *[+voice, LAB]) (see Hayes & Wilson 2008: §4.2.2 for further discussion).

Patterns that distinguish between identical and non-identical segments are an interesting test case for this model, because the distinction between identical and non-identical segments cannot be stated with standard natural classes. Generalisations over (non-)identical entities require constraints to include variables, as discussed in Marcus *et al.* (1999) and Berent *et al.* (2012). It is possible, however, to analyse the identity effect without referencing identity directly, as the cumulative effect of specific constraints on each non-identical combination. Stating the pattern in this way makes different predictions from a model where generalisations can reference identity directly; these differences are fleshed out in the remainder of this section and in §3.2, with a toy example.

A constraint schema that refers only to distinctive feature values does not allow a co-occurrence restriction on the class of non-identical segments to be stated as such. However, the identity-effect pattern can be described with a set of constraints, as the cumulative effect of multiple constraints on each individual pair of non-identical segments. To illustrate, consider a language with just three ejectives [p' t' k'], which exhibits the identity effect, as in (5a): pairs of identical ejectives are attested, but pairs of non-identical ejectives are absent. One way that this pattern could be learned is with the six constraints in (5b), which rule out each of the six ungrammatical combinations of ejectives. The constraints assume three privative place features corresponding to the three major articulators: the lips [LABIAL], the tongue tip/blade [CORONAL] and the tongue dorsum [DORSAL] (Sagey 1986, McCarthy 1988).[2,3]

[2] If place features were treated as binary, the pattern could be stated more simply as *[+lab, +cg][−lab, +cg], *[+cor +cg][−cor +cg] and *[+dors, +cg][−dors, +cg].

[3] Note that homorganicity, like identity, cannot be represented with only distinctive features. The constraint *[αPLACE, +cg][$-\alpha$PLACE, +cg], which would rule out pairs of ejectives that don't have matching place features, is thus not available in the baseline model, which has neither alpha notation nor a place node.

(5) a. *Identity effect*

p’…p’	*p’…t’	*p’…k’
t’…t’	*t’…p’	*t’…k’
k’…k’	*k’…p’	*k’…t’

b. *Constraints for identity effect*[4]

*p’…t’	*[LAB, +cg][COR, +cg]
*t’…p’	*[COR, +cg][LAB, +cg]
*k’…p’	*[DORS, +cg][LAB, +cg]
*p’…k’	*[LAB, +cg][DORS, +cg]
*t’…k’	*[COR, +cg][DORS, +cg]
*k’…t’	*[DORS, +cg][COR, +cg]

A general constraint against pairs of ejectives, *[+cg][+cg], is not compatible with the pattern in (5a), because [+constricted glottis] sequences occur on the surface. Since identical segments don’t form a natural class which is definable with distinctive features, the learner has no way of noticing that the attested combinations of [+constricted glottis] segments form a systematic class, and that all other pairs of [+constricted glottis] … [+constricted glottis] segments are absent. In the baseline model, the six constraints in (5b) cannot be condensed into one or more general constraints.

The piecemeal account of the identity effect in (5b) makes two predictions, stated here and expanded on in the remainder of this section. First, since the identity effect is formally equivalent to an arbitrary distinction between individual segment pairs, the baseline model predicts that the identity effect should be learned just as well as an arbitrary pattern, all else being equal. Second, speakers of a language with the identity effect should not extend the distinction between identical and non-identical pairs to novel segment pairs.

The first prediction is illustrated by comparing the hypothetical language on the left in (6a) below to the identity-effect pattern on the right (as in (5a)). These patterns are parallel in that three pairs of ejectives are attested and six are absent. The constraints in (6b) show that the two patterns are formally identical in a theory with no way of representing identity: both require six constraints, each of which consists of two two-feature matrices.

4 An anonymous reviewer points out that the learner may also learn general constraints like *[+cg][+cg] against pairs of ejectives, and *[COR][LAB], *[LAB][DORS], etc. against certain sequences of place feature combinations. While all of these constraints are violated by some surface forms, in principle they may be assigned some non-zero weight. Identical ejectives only violate *[+cg][+cg], while forms with non-identical ejectives violate both *[+cg][+cg] and a constraint on place-feature combinations, and thus should accumulate greater constraint violations. This analysis is possible in principle, though whether the more general constraints are assigned a non-zero weight depends on the structure of the learning data (for example, whether identical ejectives are sparsely attested or highly overattested will affect the weight of *[+cg][+cg]).

(6) a. *Arbitrary restriction*

p'…t'	*p'…p'	*p'…k'	p'…p'	*p'…t'	*p'…k'
t'…k'	*t'…t'	*t'…p'	t'…t'	*t'…p'	*t'…k'
k'…p'	*k'…k'	*k'…t'	k'…k'	*k'…p'	*k'…t'

Identity effect (right-hand columns)

b. *Constraints for arbitrary restriction*

*p'…p' *[LAB, +cg][LAB, +cg]
*t'…t' *[COR, +cg][COR, +cg]
*k'…k' *[DORS, +cg][DORS, +cg]
*p'…k' *[LAB, +cg][DORS, +cg]
*t'…p' *[COR, +cg][LAB, +cg]
*k'…t' *[DORS, +cg][COR, +cg]

Constraints for identity effect

*p'…t' *[LAB, +cg][COR, +cg]
*t'…p' *[COR, +cg][LAB, +cg]
*k'…p' *[DORS, +cg][LAB, +cg]
*p'…k' *[LAB, +cg][DORS, +cg]
*t'…k' *[COR, +cg][DORS, +cg]
*k'…t' *[DORS, +cg][COR, +cg]

The constraints that define the identity effect have no generality or simplicity advantage over the constraints that define an arbitrary pattern in (6). The formal parallelism between the arbitrary and identity-effect patterns predicts that the difference in acceptability between the attested and unattested forms in the two systems should be equally strong.[5]

The experiment presented in §4 tests this first prediction, explicitly comparing individuals' ability to learn the identity-effect pattern and an arbitrary pattern in an artificial grammar learning paradigm. The finding is that the identity-effect pattern is indeed learned better, suggesting that the correct model of phonotactics and phonotactic learning bestows some advantage to the identical/non-identical distinction.

The second prediction of a piecemeal analysis of the identity effect in the baseline model is that a distinction between identical and non-identical segments should not extend to any novel segment pairs. To illustrate, consider a language, similar to that in (5a), with three ejectives [p' t' k'] and the identity effect, but in which pairs of identical labial ejectives are absent. The pattern would be as in (7).

[5] This prediction holds when all other properties of the two languages are the same. If the relative distribution of other segments is such that attested/unattested pairs of ejectives fall into broader natural classes, then the patterns may be learned differently (e.g. if labials are not followed by coronals in a language, *p'… t' will fall under a more general constraint than *[LAB, +cg][COR, +cg]). Similarly, if the arbitrary pattern in (6a) were modified such that only two pairs of ejectives were attested, the pattern would no longer be formally equivalent to the identity effect.

(7) *Identity-effect pattern with missing* [p’…p’]

attested	*unattested* (*gap*)	*unattested* (*non-identical*)	
t’…t’	p’…p’	p’…t’	p’…k’
k’…k’		t’…p’	t’…k’
		k’…p’	k’…t’

Given the pattern in (7), constraints against individual combinations of ejectives, including pairs of labial ejectives, are learned, as shown in (8). All pairs of ejectives involving a labial ejective can be ruled out with two constraints; the remaining absent combinations are ruled out by further constraints on those individual pairs.

(8) *Constraints for identity-effect pattern with missing* [p’…p’]

*p’…p’	*p’…t’	*p’…k’	*[LAB, +cg][+cg]
*p’…p’	*t’…p’	*k’…p’	*[+cg][LAB, +cg]
*t’…k’			*[COR, +cg][DORS, +cg]
*k’…t’			*[DORS, +cg][COR, +cg]

If speakers of a language with the grammar in (8) are asked to judge the acceptability of nonce words with two ejectives, they are predicted not to distinguish between the unattested pair of identical labial ejectives and the unattested pairs of non-identical ejectives. Without constraints that refer explicitly to identical or non-identical classes of segments, no generalisation can be made for nonce pairs of identical and non-identical segments. A nonce word with [p’…p’] will not be treated as belonging to the broader class of identical ejective pairs, which are generally attested, but instead will be grouped with the other unattested non-identical ejective pairs. In fact, as will be seen in the simulations in § 3.3 below, pairs of labial ejectives are predicted to be judged to be more ill-formed than pairs of non-identical ejectives, because they are penalised both by more constraints and by more general constraints.

Many studies have found, through a variety of tasks, that speakers generalise broad patterns in their lexicon to nonce words, such that certain nonce forms emerge as more acceptable than others (Frisch & Zawaydeh 2001, Moreton 2002, Rose & King 2007, Albright 2009); cf. the well-known distinction between absent, but possible, *blick* and absent, but impossible, *bnick* in English (Halle 1962, Chomsky & Halle 1968: 380). Some evidence that generalisations formed over identical segments extend to nonce words comes from a study by Berent *et al.* (2002). Berent *et al.* investigated Hebrew speakers’ representation of a restriction on identical consonants by asking for well-formedness judgements on nonce forms containing identical pairs of consonants not native to Hebrew. Native Hebrew verbs exhibit a pattern common to many Semitic languages where identical consonants are attested at the right edge of a base form but not the left edge, e.g. [simem] ‘drug (PAST 3SG MASC)’ is an attested form, with the pattern $C_1C_2C_2$, but $C_1C_1C_2$ forms such as *[sisem] are unattested.

To test whether Hebrew speakers learn a generalisation over all identical consonants, as opposed to restrictions on individual pairs of identical consonants, Berent *et al.* constructed nonce words with identical pairs of the non-native consonants [ʧ], [ʤ] and [θ] in different positions. If Hebrew speakers' grammar references identity explicitly, then the generalisation that identical consonant pairs are well-formed at the left edge but not the right edge should extend to the novel segments. If generalisations based on the broad distinction between identical and non-identical consonants are unavailable to the learner, then identical pairs of these novel segments should be judged equally well-formed at the right edge and at the left edge. Hebrew speakers' judgements support the identity-based generalisation: forms like [kiθeθ] were given better well-formedness ratings than forms like [θiθek]. The artificial grammar experiment in §5 also finds that a distinction between identical and non-identical segments generalises to novel pairs of identical and non-identical segments, supporting the claim that identical and non-identical segments are classes that the grammar references directly.

3.2 The copying model and predictions for identity

In the learning simulations in §3.3 below, the predictions of the baseline model outlined above are tested and compared with the predictions of a model that explicitly represents identity. This copying model follows Colavin *et al.* (2010), who show that a representation of identity in the learning data allows for a better model of the patterning of identical consonants in Amharic. They found that the copying model with identity represented as reduplication is a better fit to Amharic speakers' well-formedness judgements of nonce roots than the baseline model.

In the copying model, pairs of identical consonants are represented as a single consonant followed by a placeholder X, e.g. the root sequence [k'...k'] is represented as [k'...X]. The copying model is otherwise identical to the baseline model (i.e. the constraint-induction and weighting algorithms are the same). The placeholder X is unspecified for all the features that distinguish segments in the language, but is specified for an additional, suprasegmental feature, which I call [IDENT], which is unspecified for all other consonants. This suprasegmental feature [IDENT] is different from other distinctive features, in that it refers to a relation between segments in a string, as opposed to a property of an individual segment. It could be thought of as representing double linking of a root node, as in an autosegmental representation. As with double linking, this method of representing identical consonants results in only a single phonological feature matrix for identical segments. The representation of identical and non-identical ejectives in the two models is exemplified in Table I.

As can be seen in Table I, the representation of non-identical ejective pairs does not differ between the two models. In both models, a sequence of non-identical segments is represented as a sequence of two fully

model	representation	featural representation
baseline	k'…k' k'…p'	[DORS, −cont, −son, +cg]…[DORS, −cont, −son, +cg] [DORS, −cont, −son, +cg]…[LAB, −cont, −son, +cg]
copying	k'…X k'…p'	[DORS, −cont, −son, +cg]…[IDENT] [DORS, −cont, −son, +cg]…[LAB, −cont, −son, +cg]

Table I
Representation of identical consonant pairs in the baseline and copying models.

specified feature matrices. Identical segments, however, differ in their representation between the two models. In the baseline model, an identical pair of segments is represented, like non-identical pairs, as a sequence of two fully specified feature matrices, while in the copying model there is only one fully specified feature matrix.

With identity represented as in the copying model, our toy identity-effect language can be learned with a single constraint that rules out sequences of ejectives, *[+cg][+cg]. The arbitrary pattern still cannot be learned as a broad generalisation, and requires multiple constraints, as in the baseline model. How the identity-effect and arbitrary restriction patterns could be learned in the copying model is shown in (9).

(9) a. *Arbitrary restriction*

p'…t'	*p'…X	*p'…k'
t'…k'	*t'…X	*t'…p'
k'…p'	*k'…X	*k'…t'

Identity effect

p'…X	*p'…t'	*p'…k'
t'…X	*t'…p'	*t'…k'
k'…X	*k'…p'	*k'…t'

b. *Constraints for arbitrary restriction*

*p'…X	*[+cg][IDENT]
*t'…X	*[+cg][IDENT]
*k'…X	*[+cg][IDENT]
*p'…k'	*[LAB, +cg][DORS, +cg]
*t'…p'	*[COR, +cg][LAB, +cg]
*k'…t'	*[DORS, +cg][COR, +cg]

Constraints for identity effect

*p'…t'	*[+cg][+cg]
*p'…k'	*[+cg][+cg]
*t'…p'	*[+cg][+cg]
*t'…k'	*[+cg][+cg]
*k'…p'	*[+cg][+cg]
*k'…t'	*[+cg][+cg]

As can be seen in (9b), for the arbitrary restriction in the copying model, a single constraint, *[+cg][IDENT], penalises all pairs of identical ejectives.

The other three absent pairs of ejectives, however, do not form a class distinct from the attested ejective pairs, and must be ruled out with three additional constraints on those individual ejective pairs, as in the baseline model. For the identity effect, in the copying model a single constraint, *[+cg][+cg], rules out all six unattested pairs of ejectives. This constraint does not penalise pairs of identical ejectives, because they are not represented as sequences of two [+constricted glottis] segments. In the copying model, then, the arbitrary and identity-effect patterns are not formally equivalent. Rather, the identity effect can be learned with a single, general constraint, while the arbitrary pattern requires multiple, more specific constraints. As will be shown in the simulations in the next section, more general constraints receive higher weights, and thus unattested sequences that fall into a broad class are penalised more than unattested sequences that must be ruled out by more narrow constraints. Because the copying model allows a single generalisation to be made over all non-identical segment pairs, it allows for a formal distinction between the identity effect and the arbitrary pattern. The simulations in §3.3 below show that the difference in acceptability between attested and unattested pairs in the identity-effect pattern is larger than in the arbitrary pattern in the copying model, but not the baseline model.

The copying model also differs from the baseline model in the predictions about an identity-effect system with a gap for one identical segment pair, as in (7) above, where [t'…t'] and [k'…k'] are attested, but non-identical pairs and [p'…p'] are unattested. The constraints needed to capture this pattern in the baseline model (repeated from (8)) and the copying model are presented in (10). In the baseline model, the absence of [p'…p'] allows for a broader generalisation, i.e. that labial ejectives cannot occur with any other ejective, captured by the two constraints *[LAB, +cg][+cg] and *[+cg][LAB, +cg]; the other absent combinations are again ruled out by specific constraints on those combinations. In the copying model, a single constraint cannot rule out identical and non-identical pairs of ejectives, since identical pairs are not represented as segmental sequences. Instead, all unattested non-identical ejective pairs can be ruled out with the single constraint *[+cg][+cg], and an additional, specific constraint, *[+cg, LAB][IDENT], is needed to rule out the gap.

(10) *Identity-effect pattern with missing* [p'…p']

a. *Baseline model*

*p'…p'	*p'…t'	*p'…k'	*[LAB, +cg][+cg]
*p'…p'	*t'…p'	*k'…p'	*[+cg][LAB, +cg]
*t'…k'			*[COR, +cg][DORS, +cg]
*k'…t'			*[DORS, +cg][COR, +cg]

b. *Copying model*

*p'…t'	*p'…k'	*t'…p'	*[+cg][+cg]
*t'…k'	*k'…p'	*k'…t'	
*p'…X			*[+cg, LAB][IDENT]

The crucial difference between the two models is that in the baseline model, restrictions on pairs of identical labials and non-identical pairs with a single labial can be penalised with a single constraint (e.g. *[LAB, +cg][+cg]), while in the copying model separate constraints are needed for identical and non-identical pairs. In the baseline model, grouping identical and non-identical pairs of unattested ejectives means that the [p'…p'] gap violates a more general constraint than the unattested non-identical ejectives [t'…k'] and [k'…t'] (*[LAB, +cg][+cg] refers to only three features; *[COR, +cg][DORS, +cg] and *[DORS, +cg][COR, +cg] each refer to four features). In the copying model, however, the identical pair is penalised by a more specific constraint than the non-identical pairs. As will be shown in the simulations to follow, the preference for more general constraints means that the baseline model predicts that pairs of identical labial ejectives will be more dispreferred than pairs of non-identical ejectives, while the copying model predicts the opposite.

The following section presents learning simulations that demonstrate the different predictions of the baseline model and the copying model with respect to (i) the distinction between the identity-effect pattern and an arbitrary pattern, and (ii) the extension of the identity pattern to an unattested gap.

3.3 Learning simulations: modelling the identity effect

A series of learning simulations show the predictions outlined schematically above for the baseline and copying models. The data for the simulations were based on a list of lexical roots from Chol, a Mayan language exhibiting the identity effect on ejectives (Warkentin & Brend 1974, Gallagher & Coon 2009). The consonantal inventory of Chol is given in (11) (the language has six vowels: [a e o i ɨ u]).

(11)

	labial	alveolar	palato-alveolar	palatal	velar	glottal
implosive	ɓ					
voiceless stop	p	tʲ ts	tʃ		k	ʔ
ejective	p'	tʲ' ts'	tʃ'		k'	
fricative		s	ʃ			h
affricate						
nasal	m			ɲ		
approximant	w	l		j		

Ejectives occur only in roots in Chol, which are predominantly CVC, and are restricted from co-occurring in pairs unless completely identical. A list of all 843 roots extracted from Aulie & Aulie (1978) contains 13 roots with two identical ejectives (12a), but none with two non-identical ejectives (12b).

(12) a.

pʼipʼ	‘wild’	ʧʼaʧʼ	‘bush’	kʼakʼ	‘woody vine’
tʲʼotʲʼ	‘snail’	ʧʼɨʧʼ	‘absorb’	kʼɨkʼ	‘flame’
tsʼaʰtsʼ	‘soak’	ʧʼiʧʼ	‘blood’	kʼokʼ	‘healthy’
tsʼuʰtsʼ	‘kiss’	ʧʼoʧʼ	‘throat’	kʼukʼ	‘plumage’
		ʧʼuʧʼ	(type of tree)		

b.

*pʼ…tʲʼ	*tʲʼ…pʼ	*tsʼ…pʼ	*ʧʼ…pʼ	*kʼ…pʼ
*pʼ…tsʼ	*tʲʼ…tsʼ	*tsʼ…tʲʼ	*ʧʼ…tʲʼ	*kʼ…tʲʼ
*pʼ…ʧʼ	*tʲʼ…ʧʼ	*tsʼ…ʧʼ	*ʧʼ…tsʼ	*kʼ…tsʼ
*pʼ…kʼ	*tʲʼ…kʼ	*tsʼ…kʼ	*ʧʼ…kʼ	*kʼ…ʧʼ

The learning simulations were conducted with the UCLA phonotactic learner, the inductive model proposed and tested in Hayes & Wilson (2008).[6] Six different feature simulations were run, varying both the structure of the learning data and the representations (i.e. the baseline or copying model). The learning data were either the unmodified list of Chol roots, showing the identity effect, or were modified either to show an arbitrary pattern or to exclude one of the attested pairs of identical ejectives.[7] In the arbitrary pattern, each pair of identical ejectives in the root list was changed to a different combination of non-identical ejectives. The result is a pattern where five of the 25 possible combinations of the five ejectives are attested, and 20 are unattested, as in the identity pattern. The arbitrary identity patterns here are thus parallel to those sketched schematically in §3.1; the learning data is identical in every way, except that the five attested pairs of ejectives are either all identical, or not. The second modification was to remove the root with identical labial ejectives from the learning data, so there was a gap in the identity-effect pattern: all attested pairs of ejectives are identical, but both non-identical pairs and one identical pair are unattested. The three patterns are shown in Table II, illustrating just the attested pairs of ejectives in each set of learning data; pairs of ejectives not listed are entirely absent from the learning set. The other roots (those not containing two ejectives) in the learning data were the same for all three data sets.

Simulations on each of these data structures were run twice, once without a representation of identity (the baseline model) and once with identity represented with a placeholder (the copying model). Additionally, only the consonantal root pairs were given to the model; the vowels were removed. This step was taken only for expediency, as the learner is capable of making generalisations over non-adjacent sequences, via

[6] The software is available (April 2013) at http://www.linguistics.ucla.edu/people/hayes/phonotactics/.

[7] The list of roots was repeated four times, so that the learning data had over 3000 tokens (843 × 4 = 3372). The UCLA phonotactic learner does not run reliably with fewer than 3000 tokens.

unmodified: identity		modified: arbitrary		modified: gap	
p'…p'	1	p'…tʲ'	1	tʲ'…tʲ'	1
tʲ'…tʲ'	1	tʲ'… tʃ'	1	ts'…ts'	2
ts'…ts'	2	ts'…p'	2	tʃ'… tʃ'	5
tʃ'… tʃ'	5	tʃ'…k'	5	k'…k'	4
k'…k'	4	k'…ts'	4		

Table II

Attested ejective pairs in three sets of learning data given as input to the simulations, and the number of roots with each pair (out of a total of 843).

	son	cont	lat	nas	strid	voice	cg	LAB	COR	ant	DORS	GLOT
ɓ	−	−				+		+				
p	−	−				−	−	+				
p'	−	−				−	+	+				
m	+	−		+				+				
w	+	+		−				+				
tʲ	−	−			−		−		+	+		
tʲ'	−	−			−		+		+	+		
ts	−	−			+		−		+	+		
ts'	−	−			+		+		+	+		
s	−	+			+				+	+		
tʃ	−	−			+		−		+	−		
tʃ'	−	−			+		+		+	−		
ʃ	−	+			+				+	−		
ɲ	+	−	−	+	−				+	−		
l	+	+	+	−	−				+	+		
j	+	+	−	−	−				+	−		
k	−	−					−				+	
k'	−	−					+				+	
ʔ	−	−										+
h	−	+										+

Table III

Distinctive feature chart for Chol consonants, given to the model as input to the learning simulations. Blank cells indicate that a segment is underspecified for the given feature.

n-grams or tier projection (Hayes & Wilson 2008: §6). The feature set for Chol consonants given to the model is shown in Table III.[8] All models were run to learn 200 constraints, though the same basic patterns emerge with fewer constraints.

[8] The implosive labial and glottal stops are not specified as [+constricted glottis], because these two segments do not pattern with the ejectives with respect to

The learner assigns a score to each form in the testing data, which for all simulations was the full set of attested forms in the learning data as well as all ejective pairs. The score for a given form is computed by summing the weights of the constraint violations, similar to other weighted constraint models like Harmonic Grammar (Pater *et al*. 2007, Coetzee & Pater 2008, Pater 2009, McCarthy & Pater to appear). Constraint weights and violations are positive, and thus scores have a lower bound of 0 and no upper bound. Higher scores indicate greater dispreference; a form with a score of 0 violates no constraints and is grammatically perfect. To assess each model, the scores of attested and unattested pairs of ejectives were compared. The learner is non-deterministic, meaning that it gives slightly different results on each run with the same data and settings. To get a measure of the variability between runs of the same model, each model was run five times; the tables below report the average across all five runs.

3.3.1 *Prediction 1: comparing the identity and arbitrary patterns.* The baseline and copying models make different predictions about the arbitrary and identity-effect patterns. In the baseline model, the difference between attested and unattested pairs of ejectives is roughly equivalent for both the arbitrary and identity patterns, while for the copying model the difference is bigger for the identity pattern than the arbitrary pattern. Table IV shows the average scores assigned by the baseline model to attested and unattested ejective pairs in the identity pattern and the arbitrary pattern. For both patterns, attested forms have lower scores than unattested forms. The difference between attested and unattested forms is slightly bigger for the arbitrary pattern (2·99) than for the identity pattern (2·29).

	attested	unattested
identity	1·20 (0·06)	3·49 (0·04)
arbitrary	1·31 (0·19)	4·30 (0·16)

Table IV

Scores for attested and unattested ejective pairs in the identity and arbitrary patterns, as assessed by the baseline model. Scores are averaged across five runs. One standard deviation is given in parentheses.

Table V shows the average scores for attested and unattested ejective pairs in the identity and arbitrary patterns for the copying model. Like the baseline model, the copying model assigns higher scores to unattested forms than to attested forms for both patterns. Unlike the baseline model, in the copying model the difference between attested and unattested forms is twice as large for the identity pattern (5·08) as for the arbitrary pattern (2·46).

co-occurrence restrictions. Instead, the implosive is identified as a voiced stop and glottal stop as a plain stop at the glottal place of articulation.

	attested	unattested
identity	0·66 (0·08)	5·74 (0·09)
arbitrary	1·13 (0·01)	3·59 (0·03)

Table V

Scores for attested and unattested ejective pairs in the identity and arbitrary patterns, as assessed by the copying model. Scores are averaged across five runs. One standard deviation is given in parentheses.

	constraint	weight	penalised ejective pairs
(a)	*[−ant][+ant]	2·872	*ʧ'…ts'
	*[+ant][−ant]	2·851	*ts'…ʧ'
	*[+cg, LAB][+cg, COR]	2·119	*p'…tʲ' *p'…ts' *p'…ʧ'
	*[+cg, COR][+cg, DORS]	2·081	*tʲ'…k' *ts'…k' *ʧ'…k'
	*[+cg, DORS][+cg, COR]	1·914	*k'…tʲ' *k'…ts' *k'…ʧ'
	*[+cg, LAB][+cg, DORS]	1·886	*p'…k'
	*[+cg, −strid][+cg, +strid]	1·517	*tʲ'…ts' *tʲ'…ʧ'
	*[+cg, −strid][DORS]	1·492	*tʲ'…k'
	*[−cont, +strid][+cg, LAB]	1·363	*ts'…p' *ʧ'…p'
	*[+cg, −strid][LAB]	1·228	*tʲ'…p'
(b)	*[−ant][+ant]	2·831	*ʧ'…ts'
	*[+ant][−ant]	2·479	*ts'…ʧ'
	*[+cg, −strid][DORS]	2·386	*tʲ'…k'
	*[+cg, COR][−cont, +ant]	2·226	*tʲ'…ts' *ʧ'…ts' *ts'…ts'
	*[DORS][+cg, DORS]	2·109	*k'…k'
	*[+cg, +ant][+cg, DORS]	1·955	*ts'…k'
	*[+cg, LAB][+cg, +strid]	1·790	*p'…ts' *p'…ʧ'
	*[+cg, LAB][+cg, DORS]	1·777	*p'…k'
	*[−cont, +strid][+cg, COR]	1·548	*ts'…tʲ' *ts'…ts' *ts'…ʧ' *ʧ'…tʲ' *ʧ'…ts' *ʧ'…ʧ'
	*[−son, LAB][+cg, LAB]	1·442	*p'…p'

Table VI

Ten highest-weighted constraints penalising some pair of ejectives in the baseline model: (a) identity pattern; (b) arbitrary pattern.

The baseline models of the identity and arbitrary patterns are illustrated in Table VI, which show the ten highest-weighted constraints that penalise some combination of ejectives. Some of these constraints rule out only combinations of ejectives and look like the constraints sketched in

	constraint	weight	penalised ejective pairs
(a)	*[+cg][+cg]	3·389	*all non-identical ejectives*
	*[−ant][+ant]	2·018	*ʧ’…ts’
	*[+ant][−ant]	2·004	*ts’…ʧ’
	*[+cg, −strid][DORS]	1·733	*tʲ’…k’
	*[DORS][−voi]	1·410	*k’…p’
	*[+cg, COR][−cont, +ant]	1·382	*tʲ’…ts’ *ʧ’…ts’
	*[−ant][+cg, LAB]	1·281	*ʧ’…p’
	*[+cg, LAB][IDENT]	1·204	*p’…p’
	*[+cg, −strid][IDENT]	1·204	*tʲ’…tʲ’
	*[−cont, +ant][−voi]	1·016	*ts’…p’
(b)	*[+cg][IDENT]	2·824	*all identical ejectives*
	*[+ant][−ant]	2·141	*ts’…ʧ’
	*[+cg, LAB][+cg, +strid]	2·021	*p’…ts’ *p’…ʧ’
	*[+cg, −strid][DORS]	1·960	*ts’…k’ *ʧ’…k’
	*[−ant][+ant]	1·871	*ʧ’…ts’
	*[+cg, LAB][+cg, DORS]	1·781	*p’…k’
	*[+cg, +ant][+cg, DORS]	1·759	*ts’…k’
	*[+cg, COR][−cont, +ant]	1·537	*tʲ’…ts’ *ʧ’…ts’
	*[+cg, DORS][+cg, −ant]	1·406	*k’…ʧ’
	*[DORS][+cg, LAB]	1·368	*k’…p’

Table VII

Ten highest-weighted constraints penalising some pair of ejectives in the copying model: (a) identity pattern; (b) arbitrary pattern.

the previous subsections, like *[+cg, LAB][+cg, COR]. Other constraints rule out combinations of ejectives as well as combinations of non-ejectives, like *[−ant][+ant]. The highest-weighted constraints in both models are constraints against disagreeing anteriority values, as pairs of stridents that disagree in anteriority are categorically absent in Chol. What is of interest in these models is that for both the identity (a) and the arbitrary (b) patterns, the unattested combinations of ejectives are ruled out by a variety of constraints on some subset of the unattested pairs.

The copying models for the identity and arbitrary patterns are shown in Table VII, again showing the ten highest-weighted constraints that penalise some combination of ejectives. For the identity pattern, the copying model learns a high-weighted, general constraint that rules out pairs of non-identical ejectives. The constraint *[+cg][+cg] represents this generalisation, which can be learned in the copying model because pairs of identical ejectives are not represented as sequences, and therefore do not constitute exceptions to this constraint in the training data. Additional, lower-weighted constraints further penalise subsets of the unattested, non-identical ejective pairs, as in the baseline model. This redundancy is an inherent part of the UCLA phonotactic learner, which assigns some weight to all constraints that allow for a tighter fit to the learning data. The copying model of the identity pattern also learns low weights for constraints against pairs of labial ejectives [p'…p'] and pairs of alveolar stop ejectives [tʲ'…tʲ'], which, despite being attested, are very infrequent in the learning data. In the copying model of the arbitrary pattern in Table VIIb, the unattested ejective combinations are not penalised by any single constraint. Rather, they are ruled out by constraints on subsets of the unattested combinations, or individual combinations, as in the baseline model.

The simulations reported in this subsection illustrate that the identity-effect pattern is formally equivalent to an arbitrary pattern, unless identity is represented explicitly. The baseline model learns the identity and arbitrary patterns equally well, whereas the copying model learns a stronger distinction between attested and unattested pairs for the identity pattern.

	attested identical	unattested [p'…p']	non-identical
baseline	0·93 (0·07)	8·67 (0·29)	3·67 (0·07)
copying	0·66 (0·09)	2·41 (0·13)	5·76 (0·12)

Table VIII

Scores for attested ejective pairs and unattested identical and non-identical ejective pairs, as assessed by the baseline model and the copying model. Scores are averaged across five runs. One standard deviation is given in parentheses.

3.3.2 *Prediction 2: generalisation of the identity pattern.* The second set of simulations compared the baseline and copying models on the learning data with a gap for identical labial ejectives. Table VIII shows the results of both models. Here, the scores are shown for attested identical ejectives, the pair of identical labial ejectives that is unattested in the learning data and the unattested non-identical ejectives. In both models, unattested non-identical ejectives are given higher scores than pairs of attested identical ejectives, though the difference is bigger for the copying model than the baseline model (as found for the simulations in the previous section). Crucially, the models differ with respect to the unattested pair of two labial ejectives. In both models, the unattested identical labial ejectives have a higher score than the attested pairs of identical ejectives. In the baseline model, however, this pair is more ill-formed – has a higher score – than the other unattested non-identical ejectives, while in the copying model the identical labial ejective pair is less ill-formed – has a lower score – than the non-identical ejectives. That is, in the copying model the grammaticality of identical ejectives in the learning data is

	constraint	weight	penalised ejective pairs
(a)	*[−ant][+ant]	2·867	*tʃ’…ts’
	*[+ant][−ant]	2·865	*ts’…tʃ’
	*[+cg, LAB][+cg]	2·249	*p’…p’ *p’…tʲ’ *p’…ts’ *p’…tʃ’ *p’…k’
	*[+cg, COR][+cg, DORS]	2·168	*tʲ’…k’ *ts’…k’ *tʃ’…k’
	*[+cg, DORS][+cg, COR]	2·006	*k’…tʲ’ *k’…ts’ *k’…tʃ’
	*[+cg, −strid][+cg, +strid]	1·512	*tʲ’…ts’ *tʲ’…tʃ’
	*[+cg, −strid][DORS]	1·481	*tʲ’…k’
	*[+cg, −strid][LAB]	1·453	*tʲ’…p’
	*[+cg, +strid][+cg, −strid]	1·367	*ts’…tʲ’ *tʃ’…tʲ’
	*[−son, LAB][+cg, LAB]	1·346	*p’…p’
	*[+cg, LAB][−cont, LAB]	1·008	*p’…p’
	*[LAB][−voi]	0·957	*p’…p’
	*[+cg][−voi]	0·847	*p’…p’ *tʲ’…p’ *ts’…p’ *tʃ’…p’ *k’…p’
	*[+cg, LAB][−son, LAB]	0·825	*p’…p’
	*[−son, −cont][+cg, LAB]	0·779	*p’…p’ *tʲ’…p’ *ts’…p’ *tʃ’…p’ *k’…p’
	*[+cg][+cg, LAB]	0·695	*p’…p’ *tʲ’…p’ *ts’…p’ *tʃ’…p’ *k’…p’
	*[+cg][+cg]	0·231	*all ejectives*
(b)	*[+cg][+cg]	3·248	*all non-identical ejectives*
	*[+ant][−ant]	2·136	*ts’…tʃ’
	*[−ant][+ant]	1·986	*tʃ’…ts’
	*[+cg, LAB][IDENT]	1·931	*p’…p’
	*[+cg, −strid][DORS]	1·487	*tʲ’…k’
	*[+cg, −strid][LAB]	1·288	*tʲ’…p’
	*[+cg, COR][−cont, +ant]	1·131	*tʲ’…ts’ *tʃ’…ts’
	*[−ant][+cg, LAB]	1·008	*tʃ’…p’
	*[+cg, −strid][IDENT]	0·983	*tʲ’…tʲ’
	*[LAB][+cg, −strid]	0·896	*p’…tʲ’
	*[−voi][IDENT]	0·483	*p’…p’
	*[wordboundary][+cg, LAB]	0·327	*p’…p’ *p’…tʲ’ *p’…ts’ *p’…tʃ’ *p’…k’
	*[+cg][IDENT]	0·211	*all identical ejectives*

Table IX

Ten highest-weighted constraints penalising some pair of ejectives for the gap pattern in (a) the baseline model; (b) the copying model. Lower weighted constraints penalising labial ejectives are below the dashed line.

extended to an unattested identical pair, while in the baseline model it is not.

The constraints that the two models learn are illustrated in Table IX, which shows the ten highest-weighted constraints that penalise some combination of ejectives. Additionally, all constraints that penalise labial ejective pairs are shown (below the dashed line). The main difference between the two models is that, in the baseline model (a), the absence of identical pairs of labial ejectives allows the model to learn a high weight for a general constraint against all combinations of a labial ejective followed by some other ejective, *[+cg, LAB][+cg]. In the copying model (b), this general constraint is not learned, because identical ejectives are not represented as sequences of ejectives, and thus pairs of identical segments don't fall within the scope of broader generalisations that hold of non-identical segment pairs. Instead, in the copying model pairs of labial ejectives are penalised by a more specific, and lower-weighted, constraint specifically against identical labial ejectives *[+cg, LAB][IDENT]. While both models learn several lower-weighted constraints that penalise labial ejectives, in the baseline model these constraints are more general and higher-weighted than in the copying model.

The simulations reported in this subsection illustrate that the identity-effect pattern is extended to novel, unattested ejective pairs in the copying model but not the baseline model.

3.4 Summary

The section has shown that the predictions of the baseline model, which lacks an explicit representation of identity, differ from those of the copying model, which represents identity. In the baseline model, the distinction between attested and unattested ejectives is learned equally well for an identity-effect pattern and an arbitrary pattern, while in the copying model the identity-effect pattern is learned better (the difference between attested and unattested ejective pairs is greater). Furthermore, in the copying model the distinction between identical and non-identical ejectives is extended to a novel pair of identical ejectives, while in the baseline model it is not. In sum, a model in which identity is represented allows for a greater distinction between identical and non-identical segments to be learned, for the distinction to be learned as a broad, systematic generalisation, and for generalisation of the identity effect to unattested segment pairs.

The source of the different predictions of the baseline and copying models is in the generality of the constraints they can learn. Because of both the accuracy and generality heuristics in the constraint-weighting algorithm, a pattern that can be stated with a single constraint will be learned with a higher weight than one that can only be stated with several constraints. A constraint against all non-identical ejectives is preferred on accuracy, because there is more data to support it than a constraint such as just, say, *k'…p', and it is more general because it refers to fewer features. This feature of the learning algorithm is desirable, as it has been found that formal properties like scope and simplicity influence learning in artificial grammar experiments (Moreton 2008, 2012, Pater & Moreton 2012). The experiments in the next section show that an identity-based pattern is indeed learned better than an arbitrary pattern, and thus warrants a simpler statement, as in the copying model.

4 Experiment 1: learning the identity effect in an artificial grammar paradigm

An artificial grammar learning experiment was designed to test how well a pattern based on a distinction between identical and non-identical consonants is learned relative to a pattern based on an arbitrary distinction among pairs of consonants. Participants were taught a simple language game in which they were trained on an alternation between voiced and voiceless stops. In one condition, all stops underwent the alternation except for those that co-occurred with an identical stop (i.e. stimuli with [b…b, d…d, g…g] pairs patterned differently from stimuli with [b…d, b…g, d…b, d…g, g…b, g…d] pairs). In a second condition, an arbitrary distinction was made between consonant pairs, such that stimuli with three arbitrarily chosen pairs of consonants did not exhibit an alternation, while stimuli with the remaining six pairs of consonants did (i.e. stimuli with [b…d, d…g, g…b] pairs patterned differently from stimuli with [b…b, b…g, d…b, d…d, g…d, g…g] pairs).

If generalisations are learned at the segmental level, over individual combinations of segments, then the two conditions are formally equivalent. In each case, stimuli with six pairs of consonants exhibit an alternation, while three pairs of consonants do not. If identity is a representational primitive, however, and learners can make generalisations that refer to identity, then the pattern based on identity can be stated more simply than the arbitrary pattern.

4.1 Method

4.1.1 *Participants.* Participants were recruited via Mechanical Turk, Amazon's crowd-sourcing site,[9] and were paid \$0·60 for completing the experiment, which took about ten minutes. Participants were told that they needed to be native speakers of English to complete the experiment, and were asked in a short demographic survey at the end of the experiment what their native language was. In addition, participants were restricted to those with a 95 % approval rating on previous Mechanical Turk tasks, and an IP address within the United States.

Participants' data were excluded from analysis if (i) they stated at the end of the experiment that they were not native speakers of English (5 participants), (ii) they gave no alternating responses (13 participants) or (iii) more than 50 % of their responses were erroneous (8 participants; see §4.3 below for what classifies a response as erroneous). Once participants were removed, responses from 102 participants remained in each of the

[9] See http://www.mturk.com.

IDENTITY and SEGMENTAL conditions and 104 participants in the CONTROL condition.

4.1.2 *Stimuli.* The stimuli for the experiment were $C_1V_1C_2V_2$ nonce words, where V_1 and V_2 were all non-identical pairings of the vowels [a e i u], and C_1 and C_2 were identical and non-identical combinations of [b d g] in C_1 and [b d g p t k] in C_2. In addition, filler items were made with the same vowel patterns and all combinations of [ʧ ʃ] in C_1 and [b d g] in C_2. Stimuli were thus items like [badu], [dike], [giga], [ʧebu], etc. There was a total of 288 stimuli following this template: 216 critical items and 72 fillers.

Stimuli were made from recordings of a native Hebrew speaker, who was naive as to the purpose of the experiment, reading each stimulus item. A Hebrew speaker was chosen because the language has a canonical five-vowel inventory with no vowel reduction in unstressed positions, and a contrast between fully voiced and voiceless unaspirated stops.[10] The speaker was instructed to read the stimulus items as if they were Hebrew nonce words, with stress on the initial syllable. The words were read off a print-out, and each one was repeated twice. The first and last item on each page were discarded, to avoid stimuli with prosodic effects associated with the beginning/end of a list. The second repetition of each word was used for the experimental stimuli, except for a small number of cases where the final vowel was either devoiced or very quiet, in which case the initial repetition was used. The stimuli were not modified in any way.

There were two conditions, exhibiting one of two patterns. In the IDENTITY condition, identical and non-identical pairs of consonants patterned differently: a prompt with identical voiced stops was paired with an identical response (e.g. prompt [babu], response [babu]), while a prompt with non-identical voiced stops was paired with a response with the second stop devoiced (e.g. prompt [badu], response [batu]; prompt [bagu], response [baku]). In the SEGMENTAL condition, prompts with labial … coronal, coronal … dorsal and dorsal … labial pairs of voiced stops were paired with an identical response (e.g. prompt [badu], response [badu]), while prompts with all other combinations of voiced stops were

[10] Hebrew was used for creating the stimuli instead of English in order to avoid any possible confounds introduced by the realisation of the voicing contrast in English. English exhibits an aspiration contrast in initial position, but a voicing contrast in intervocalic unstressed position (Lisker & Abramson 1967, Abramson & Lisker 1970). By contrast, Hebrew stops show a voiced *vs.* voiceless opposition in both initial and intervocalic position. It is possible, however, that Hebrew voicing categories are misperceived by English-speaking participants, specifically in that initial voiceless unaspirated stops are categorised as voiced. Participants learned the patterns in both conditions, suggesting that they were able to accurately perceive the stimuli. In addition, stimuli were presented orthographically as well as auditorily, to minimise the possibility or effects of misperception. The Hebrew stimuli may introduce an unwanted categorisation problem, and natural English stimuli may be better for any future experiments, particularly if there is no orthographic redundancy.

IDENTITY				SEGMENTAL			
alternating		non-alternating		alternate		non-alternating	
prompt	response	prompt	response	prompt	response	prompt	response
badu	batu	babu	babu	babu	bapu	badu	badu
bagu	baku	dadu	dadu	bagu	baku	dagu	dagu
dabu	dapu	gagu	gagu	dabu	dapu	gabu	gabu
dagu	daku			dadu	datu		
gabu	gapu			gadu	gatu		
gadu	gatu			gagu	gaku		

Table X

Stimuli for the IDENTITY and SEGMENTAL experimental conditions.

paired with responses with the second stop devoiced (e.g. prompt [bagu], response [baku]). The SEGMENTAL condition shows a pattern similar to the patterns referred to as 'arbitrary' in the previous section. The two patterns are schematised in Table X, using stimuli with the [a…u] vowel pattern. Fillers were only present in testing; there were no fillers in the training trials.

In addition to the IDENTITY and SEGMENTAL conditions, a third condition, the CONTROL condition, was run. The goal of the CONTROL condition is to assess bias in participants that may account for learning differences between the identity and segmental conditions (e.g. does the English lexicon bias participants from giving responses with the [b…b] consonant pair?). In the CONTROL condition, participants were trained on 50% alternating pairs and 50% non-alternating pairs, which showed no consistent pattern. A prompt stimulus with any consonant pair could appear in training with either an alternating or a non-alternating response, and thus no generalisation could be made about when alternation occurred. For each prompt, either an alternating or a non-alternating response was chosen at random for each participant. An illustration of what training stimuli in the CONTROL condition could look like is shown in Table XI. This table shows six prompts, and how responses could be randomly assigned to two participants. For both participants, half of the responses are alternating and half are non-alternating, but the alternating/non-alternating prompts are not necessarily the same across participants (though they may be, accidentally, as for [bage] and [dedu] in Table XI), or consistent for a given consonant pair (though they may be accidentally, as for the two [d…d] prompts for participant $n+1$ in Table XI).

The testing prompts were drawn evenly from each of the three sets of prompts that are compared in the IDENTITY and SEGMENTAL conditions;

prompt	response participant n	response participant $n+1$
badu	badu	batu
beda	beta	beda
giba	gipa	giba
dedu	detu	detu
deda	deda	deta
bage	bage	bage

Table XI
Illustration of possible prompt–response pair selection for two hypothetical participants in the CONTROL condition.

seven stimuli were drawn from the set of non-alternating prompts in the IDENTITY condition ([b…b, d…d, g…g]), seven from the set of non-alternating prompts in the SEGMENTAL condition ([b…d, d…g, g…b]) and seven from the set of prompts that are alternating in both conditions ([b…g, d…b, g…d]).

4.1.3 *Equipment*. The experiment was run using Experigen, a program for running online experiments (Becker & Levine 2010). Participants were directed to the experiment, hosted on the author's university webspace, from Mechanical Turk.

4.1.4 *Procedure*. The procedure employed is similar to that in Wilson (2006). Participants were taught a language game, in which they were presented with a prompt and had to provide a response. In training, subjects were presented with two stimulus items, the prompt and the response, in the frame 'I say "X"', 'You say "Y"'. In testing, subjects were presented with a prompt and were asked to provide a response.

The experiment consisted of a training and a testing phase. In the CRITICAL training conditions, the training phase was made up of 36 prompt–response pairs, 18 drawn each from the alternating and non-alternating sets for the relevant condition. The set of 36 training pairs was repeated twice, in random order, for a total of 72 training trials. In a training trial, the participant was presented with the text 'I say "X"', with the stimulus displayed orthographically, followed by a button that the participant clicked to listen to the prompt stimulus. Once the prompt stimulus was played, a second line of text, 'You say "Y"', appeared, along with a button to listen to the response stimulus. After the response stimulus played, a 'continue' button appeared at the bottom of the page, which the participant clicked to move on to the next trial. Participants thus had to listen to both the prompt and the response stimulus before moving on. Participants were instructed to repeat the response stimulus to

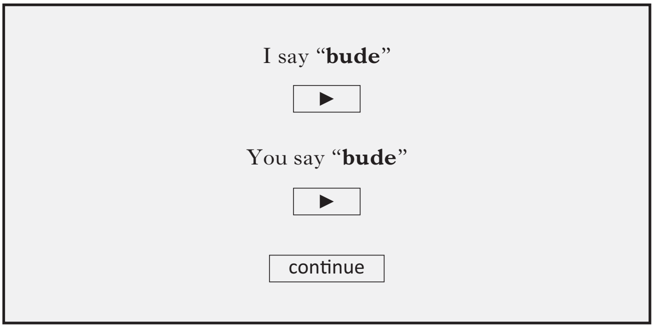


Figure 1
Screenshot of a training trial.

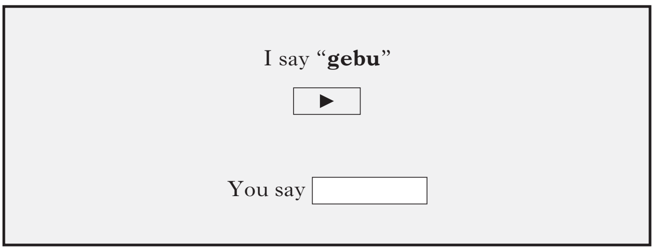


Figure 2
Screenshot of a testing trial.

help them learn the pattern, but as the experiment was run online and subjects were not recorded, it is not known what proportion of participants actually did repeat the response stimulus. A screenshot for the training phase is given in Fig. 1.

The testing phase consisted of 26 items, ten prompts each from the alternating and non-alternating sets, and six fillers. In testing, subjects were presented with the prompt stimulus in the same format as in training, the text 'I say "**x**"' followed by a button to play the stimulus. Once the prompt had been played, the text 'You say' appeared followed by a response box, where participants typed in their response. A screenshot for the testing phase is given in Fig. 2. The open response format was chosen instead of a forced choice design, in the hope that it would make the experiment more engaging. As will be mentioned in §4.3, this response format did result in needing to discard some anomalous responses, but overall participants used the orthography properly and entered either a response identical to the prompt, or one with the second stop devoiced (e.g. responses to the prompt [badu] were overwhelmingly either *badu* or *batu*).

Items from training were never repeated in testing, so while all testing items contained consonant combinations seen in training, the

IDENTITY			SEGMENTAL		
alternating prompt	alternating response	non-alternating response	alternating prompt	alternating response	non-alternating response
badu	batu	badu	babu	bapu	babu
bagu	baku	bagu	bagu	baku	bagu
dabu	dapu	dabu	dabu	dapu	dabu
dagu	daku	dagu	dadu	datu	dadu
gabu	gapu	gabu	gadu	gatu	gadu
gadu	gatu	gadu	gagu	gaku	gagu
non-alternating prompt	alternating response	non-alternating response	non-alternating prompt	alternating response	non-alternating response
babu	bapu	babu	badu	batu	badu
dadu	datu	dadu	dagu	daku	dagu
gagu	gaku	gagu	gabu	gapu	gabu

Table XII

Examples of alternating and non-alternating responses for the IDENTITY and SEGMENTAL patterns for each consonant combination, by prompt type.

particular item had not been seen before (e.g. if [bagu], [biga] and [buga] appeared in training as stimuli with the consonant pair [b…g], [bega] might appear in testing). Additionally, training and testing items were selected randomly from the set of all stimuli for each participant, so neither the set of training items nor the set of testing items were the same across participants.

4.2 Predictions

This experiment tests the first distinction between the baseline and copying models, outlined in §3 above. The baseline model predicts that participants should learn the patterns in the IDENTITY and SEGMENTAL conditions equally well, while the copying model predicts that the participants should learn the pattern in the IDENTITY condition better. Both models predict that both the IDENTITY and SEGMENTAL condition patterns should be learnable, as both models assigned a higher score to unattested segment pairs than to attested segment pairs in the simulations.

4.3 Results and analysis

For the IDENTITY and SEGMENTAL conditions, responses to each trial were coded for whether they alternated or not, and for whether the prompt triggered an alternating response in training or not. Table XII shows how

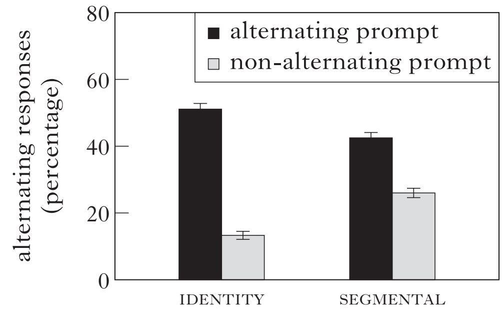


Figure 3

Proportion of alternating responses given to alternating and non-alternating prompts, by condition. Error bars indicate the standard error measure.

responses were coded for each condition, using stimuli with the [a…u] vowel pattern as illustration. Alternating responses were coded as '1' and non-alternating responses as '0'.

Some responses were removed from the analysis because the response was not classifiable as 'alternating' or 'non-alternating'. Responses were removed either because the participant gave no response (pressing the enter key in the testing phase moved on to the next trial, regardless of whether there was anything entered in the response box), or because the response differed from the prompt in place of articulation (e.g. prompt [badu], response *baku*). Responses were *not* excluded if one or both vowels did not match the prompt (e.g. *bida* instead of *bide*), or if vowels were spelled in English orthography instead of in the orthography used in the experiment (e.g. *bidoo* instead of *bidu*). 858 individual responses (11 % of the data) were excluded (261 from the IDENTITY condition, 257 from the SEGMENTAL condition and 340 from the CONTROL condition). Finally, responses to the fillers were also not analysed.

The comparison between the IDENTITY and SEGMENTAL conditions is shown graphically in Fig. 3. For both conditions, participants gave more alternating responses to alternating prompts than to non-alternating prompts. The difference between alternating and non-alternating prompts is bigger in the identity condition than the segmental condition, however. In the IDENTITY condition, participants gave alternating responses to alternating prompts more often than in the SEGMENTAL condition, and gave non-alternating responses to non-alternating prompts more often than in the SEGMENTAL condition.

A Mixed Logit Model was fit using the *lmer*() function from the lme4 package (Bates & Maechler 2010) for the R software. The dependent variable was whether the response alternated or not. The predictors in the model were condition (IDENTITY *vs.* SEGMENTAL) and prompt type (alternating or not), and the interaction between condition and prompt type. Random intercepts were included for participant and prompt, and a random slope for condition by prompt stimulus. Random slopes were not included for participant, since participants were nested within conditions; in contrast, each prompt stimulus appeared in both conditions. Significance of factors was assessed using model comparison. Removing the interaction factor resulted in a significant difference from the full model ($p < 0{\cdot}001$), showing that the full model, given in Table XIII, is a better fit to the data than a simpler model without the interaction.[11] There are significant main effects of condition and prompt type, as well as a significant interaction.

	estimate	standard error	Wald's z	p
intercept	0·74	0·06	11·89	< 0·0001
condition (segmental)	0·66	0·13	5·17	< 0·0001
prompt type (alternating)	1·73	0·09	20·27	< 0·0001
condition : type	0·55	0·16	3·35	< 0·001

Table XIII

Results of a Mixed Logit Model testing for main effects of condition (baseline IDENTITY) and prompt type (baseline non-alternating) and an interaction between condition and prompt type on alternation rate in responses.

The main effect of condition results from the greater proportion of alternating responses in the SEGMENTAL condition than in the IDENTITY condition (34% *vs*. 32%). The main effect of prompt type indicates that participants in both conditions were affected by training, giving more alternating responses to alternating prompts (51% in the IDENTITY condition and 43% in the SEGMENTAL condition) than to non-alternating prompts (13% in the IDENTITY condition and 26% in the SEGMENTAL condition). Finally, the interaction between condition and prompt type results from the larger difference between prompt types in the IDENTITY condition than in the SEGMENTAL condition.

In the CONTROL condition, participants gave an alternating response at about the same rate regardless of prompt type: 34% for stimuli drawn from the set of non-alternating prompts in the IDENTITY condition, 35% for stimuli drawn from the set of non-alternating prompts in the SEGMENTAL condition and 33% for stimuli drawn from the set of prompts that alternate in both conditions. Two Mixed Logit Models were fit to explicitly compare the CONTROL condition with the IDENTITY and SEGMENTAL conditions. One model predicted rate of alternation in responses by condition, comparing the CONTROL condition with the alternating prompts in the IDENTITY and SEGMENTAL conditions. The

[11] The command entered in the R console to run the statistical analysis was: lmer(alt_response ~ condition + alt_prompt + condition :alt_prompt + (1 | subject) + (1 + condition | prompt_stimulus)).

second model compared the CONTROL condition with the non-alternating prompts in the IDENTITY and SEGMENTAL conditions. Two analyses are necessary, because prompts in the CONTROL condition cannot be coded as alternating or non-alternating, since this was randomised by subject and stimulus item; thus there is no consistent coding schema across the three conditions. For both models, random intercepts were included for participant and prompt, and a random slope for condition by prompt stimulus. The CONTROL condition was set as the baseline for the condition factor, and a significant difference was found for both the SEGMENTAL and IDENTITY conditions, for both alternating (Table XIVa) and non-alternating (b) prompts.

		estimate	standard error	Wald's z	p
(a)	intercept	−0·81	0·12	−6·58	< 0·0001
	condition (identity)	0·88	0·18	4·80	< 0·0001
	condition (segmental)	0·42	0·18	2·40	< 0·02
(b)	intercept	−0·80	0·11	−7·18	< 0·0001
	condition (identity)	−1·40	0·18	−7·74	< 0·0001
	condition (segmental)	−0·43	0·16	−2·62	< 0·01

Table XIV

Results of a Mixed Logit Model comparing rate of alternation in the CONTROL condition to rate of alternation to (a) alternating and (b) non-alternating prompts in the IDENTITY and SEGMENTAL conditions.

In both the IDENTITY and SEGMENTAL conditions, alternating responses to non-alternating prompts were lower than in the CONTROL condition, and alternating responses to alternating prompts were higher than in the CONTROL condition, showing that the results of the IDENTITY and SEGMENTAL conditions are unambiguously due to training.

4.4 Discussion

The results of Experiment 1 show that participants learn both the identity and segmental pattern, but that they learn the identity pattern better. The better learning on the identity pattern suggests that the representational distinction between identical and non-identical segments is available to learners. While both the segmental and identity patterns can be learned as generalisations over individual pairs of segments, only the identity pattern is compatible with a broader generalisation.

The results of the experiment are consistent with the predictions of the copying model. The simulations in §3.3 showed that the copying model is able to learn general constraints on all non-identical segments, and that these constraints are more highly weighted than more specific constraints on individual segment pairs. The generality heuristic in the UCLA phonotactic learner prioritises more general constraints, and thus distinctions between attested and unattested (or frequent *vs.* infrequent) structures that fall under more general constraints are learned better.

The experiments in this paper do not explicitly test whether participants are learning generalisations that refer to whole segment identity, or to homorganicity. The identity pattern could also be described and learned as a dependency between consonants with the same place of articulation: the second stop is voiceless unless it is homorganic with the first stop. While future work should tease apart this confound, it does not interfere with the main argument in this paper. Both whole segment identity and homorganicity require explicit reference to matching, either with variables over all features (or at the root node) or over some particular features (or at the place node).

The results also show that participants are biased against giving an alternating response. Across all three training conditions, participants gave about one-third alternating responses, despite the training data consisting of 50% alternating responses. A similar effect was found in the artificial grammar learning experiments in Wilson (2006) and Coetzee (2009b). Coetzee hypothesises that learners are biased against alternation, both in artificial grammar tasks and real language learning, and will only freely generalise an alternation with very strong evidence from learning data. While interesting, this bias against alternation does not bear on the main issue of the paper, which is the distinction between identical and non-identical segments, and will not be discussed further.

5 Experiment 2: generalising the identity effect to novel segment pairs

Experiment 2 tests whether a distinction between identical and non-identical segment pairs extends to novel identical and non-identical pairs. In Experiment 1, the prompts in the training and testing stimuli consisted of the same combinations of consonants. In Experiment 2, one pair of non-identical consonants and one pair of identical consonants are omitted from the training stimuli, which otherwise show the identity effect. The testing phase then compares how well participants learn the identical/non-identical distinction for stimuli with consonant combinations they saw in training, and whether they extend this distinction to new items.

5.1 Method

5.1.1 *Participants.* Participants were recruited via Mechanical Turk, and were paid $0·60 for completing the experiment, which took about ten minutes. Responses from 44 participants were excluded, 24 because they gave zero alternating responses and 20 because 50% or more of their responses were unusable. Once participants were removed, responses from 193 participants remained for analysis, 97 from the CRITICAL condition and 96 from the CONTROL condition.

5.1.2 *Stimuli.* The stimuli were the same as in Experiment 1. Participants were trained on an identity-effect pattern, as in the IDENTITY condition in Experiment 1, except that one identical pair of segments and one non-identical pair were not included in training. In testing, prompts either contained pairs of consonants attested in the training data (old) or novel consonant pairs (new). The pattern is schematised in Table XV, using stimuli with the [a…u] vowel pattern. As in Experiment 1, fillers were only included in testing; there were no fillers in training.

TRAINING (prompt–response pair)				TESTING (prompt only)			
alternate		non-alternating		alternate		non-alternating	
prompt	response	prompt	response	old	new	old	new
badu	batu	babu	babu	badu	dagu	babu	gagu
bagu	baku	dadu	dadu	bagu		dadu	
dabu	dapu			dabu			
gabu	gapu			gabu			
gadu	gatu			gadu			

Table XV
Sample of stimuli for Experiment 2.

In addition to the CRITICAL training condition schematised in Table XV, a CONTROL condition was also run. In the CONTROL condition, prompts were randomly paired with alternating or non-alternating responses, as in the CONTROL condition for Experiment 1. In the CONTROL condition for Experiment 2, prompts with [g…g] and [d…g] consonant pairs were also omitted from training, but included in testing. The rate of alternating responses in the CONTROL condition provides a baseline for comparing the rate of alternating responses on alternating and non-alternating new and old prompt types in the CRITICAL condition.

5.1.3 *Equipment.* The equipment was the same as for Experiment 1.

5.1.4 *Procedure.* The structure of the experiment was the same as in Experiment 1. In the CRITICAL condition, there were 24 training pairs showing the pattern in Table XIII above, 12 each drawn from the alternating and non-alternating sets. The 24 training pairs were repeated twice, in random order, for a total of 48 training trials. In testing, there were eight prompt stimuli each from the alternating–old, alternating–new, non-alternating–old and non-alternating–new categories, and six fillers, for a total of 38 testing trials.

In the CONTROL condition, participants were also trained on 50% alternating and 50% non-alternating pairs, drawn from the same set of training pairs in Table XIII, but showing no consistent pattern. The structure of the training phase of the CONTROL condition was the same as for Experiment 1. The testing phase of the CONTROL condition was the same as the CRITICAL training condition.

As in Experiment 1, stimulus items were randomly selected for each participant from the full set of items, and were not repeated from training to testing.

5.2 Predictions

The results of Experiment 1 show that participants learn the identity-effect pattern, and thus participants in Experiment 2 are predicted to give more alternating responses to old prompts with non-identical consonant pairs than to old prompts with identical consonant pairs. The key question for this experiment is whether the proportion of alternating responses differs between the new non-identical consonant pair and the new identical consonant pair. If participants learn a broad generalisation about identical and non-identical consonant pairs, then they should give fewer alternating responses to prompts with identical [g…g] than to those with non-identical [d…g].

This experiment is in parallel with the simulations in §3.3.2, testing how the baseline and copying models extend an identity-based pattern to novel segments. The baseline model predicts that the identity distinction should not generalise. Rather, participants should only make generalisations based on the behaviour of individual segments or segment pairs. Given the learning data in the CRITICAL condition, participants could notice that there are no responses with [g] in C_2, i.e. the only prompts with [g] in C_2 are [b…g] prompts, which are paired with an alternating [b…k] response. This observation would lead to high rates of alternating responses to both [g…g] and [d…g] prompts in testing. Participants could also notice the rates of alternation of prompts with initial [g] or [d] in training. Prompts with initial [g] are all alternating ([g…b] and [g…d]), and thus again the prediction is made that participants should give an alternating response to [g…g] prompts. One set of prompts with initial [d] is alternating in training ([d…b]) and one set is not ([d…d]), so the prediction would be that participants should give half alternating and half non-alternating responses to [d…g] prompts in testing. The copying model predicts that participants should learn not just the distribution of individual segments or segment combinations, but also generalisations based on identity. If the copying model is right, participants should be able to generalise the distinction between identical and non-identical segments in training, and give more alternating responses to prompts with non-identical [d…g] than to those with identical [g…g].

5.3 Results and analysis

For each condition, responses were coded for whether they alternated or not, for whether the prompt was 'old' (the consonant pair was present in training) or 'new' (the consonant pair was not present in training), and for whether the consonants were identical. Table XVI shows how responses were classified, using stimuli with the [a…u] vowel pattern to illustrate.

		prompt	alternating	non-alternating
old consonant pair	identical	babu	bapu	babu
		dadu	datu	dadu
	non-identical	badu	batu	badu
		bagu	baku	bagu
		dabu	dapu	dabu
		gabu	gapu	gabu
		gadu	gatu	gadu
new consonant pair	identical	gagu	gaku	gagu
	non-identical	dagu	daku	dagu

Table XVI

Examples of alternating and non-alternating responses for new and old stimuli, by consonant combination.

As in Experiment 1, some responses were excluded from analysis because they could not be classified according to the coding scheme for the reasons described above for Experiment 1. In total, 746 individual responses (12% of the data) were excluded (378 from the CRITICAL condition and 368 from the CONTROL condition). As in Experiment 1, responses to the fillers were not analysed.

The results are shown graphically in Fig. 4. For both old and new stimuli, the rate of alternation conforms to the pattern in training in the CRITICAL condition but not the CONTROL condition. Participants gave more alternating responses to stimuli with non-identical consonants than to those with identical consonants in the CRITICAL condition but not in the CONTROL condition. Moreover, participants show this pattern for stimuli with old consonant pairs they saw in training, and also extend the pattern to stimuli with new consonant pairs.

A Mixed Logit Model was fit, predicting rate of alternation from three binary factors. The predictors compared the CONTROL and CRITICAL

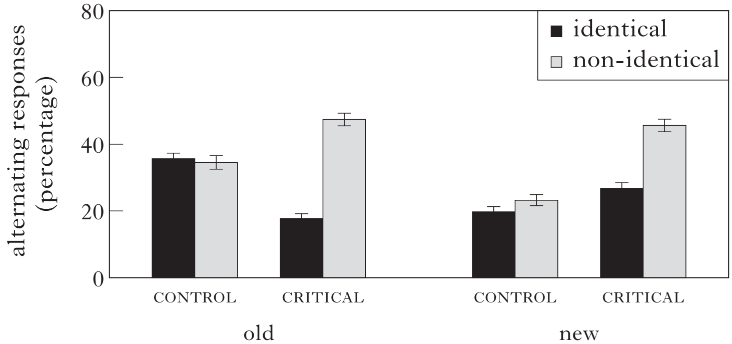


Figure 4

Proportion of alternating responses to old and new stimuli in the CRITICAL and CONTROL conditions, for stimuli with identical consonants (which alternated in CRITICAL training) and those with non-identical consonants (which did not alternate in CRITICAL training). Errors bars indicate the standard error measure.

conditions (control *vs*. critical training), stimuli with consonant pairs that appeared in training and stimuli with pairs that did not (old *vs*. new) and stimuli with identical and non-identical consonants (ident *vs*. non-ident). All two-way interactions between these factors were included, as well as the three-way interaction. Random intercepts were included for participant and prompt stimulus. Additionally, random slopes for consonant pair and prompt type by participant and a random slope for training by prompt stimulus were also included. Neither the three-way interaction between factors nor the interaction between consonant pair and prompt type was significant (as assessed by model comparison). The final model is given in Table XVII.[12]

The main effect of training is found because participants gave more alternating responses in the CRITICAL condition (34 %) than in the CONTROL condition (28 %). Participants also gave more alternating responses to prompts with non-identical consonants (38 %) than identical consonants (25 %), in conformance with the pattern in the CRITICAL condition. The main effect of old *vs*. new consonant pair reflects the greater proportion of alternating responses to stimuli with old consonant pairs (34 %) than new consonant pairs (29 %). The key result is the significant interaction between training and ident *vs*. non-ident, which results from the greater effect of prompt type (identical or non-identical consonants) on responses in the CRITICAL condition than in the CONTROL condition. Participants gave more alternating responses to stimuli with pairs of old

[12] The command entered in the R console to run the statistical analysis was: lmer (alt_response ~ training + cons_pair + prompt_type + training:cons_pair + training: prompt_type + (1 + cons_pair + prompt_type + cons_pair:prompt_type|subject) + (1 + training|stimulus).

	estimate	standard error	Wald's z	p
intercept	−1·42	0·11	−12·57	< 0·0001
training (control)	−0·52	0·23	−2·27	< 0·03
ident *vs*. non-ident (non-ident)	0·76	0·12	6·11	< 0·0001
old *vs*. new (old)	1·11	0·18	6·05	< 0·0001
training : ident *vs*. non-ident	−1·45	0·26	−5·60	< 0·0001
training : old *vs*. new	1·53	0·37	4·12	< 0·0001

Table XVII

Results of Mixed Logit Model testing for main effects of training (baseline CRITICAL condition) and consonant pair (baseline new) and an interaction between training and consonant pair on conforming responses.

and new non-identical consonants than to pairs of old and new identical consonants in the CRITICAL condition (46% for non-ident *vs*. 22% for ident), but showed no such effect in the CONTROL condition (28% for non-ident *vs*. 29% for ident). Finally, the significant interaction between training and old *vs*. new arises because participants in the CRITICAL condition gave alternating responses at similar rates to stimuli with both old and new consonant pairs (32% for old *vs*. 36% for new), while participants in the CONTROL condition gave alternating responses much more often to stimuli with old consonant pairs than to those with new consonant pairs (35% for old *vs*. 21% for new).

5.4 Discussion

Participants learn a pattern that distinguishes between identical and non-identical consonant pairs, and extend this pattern to novel segment pairs. The extension of the identity-effect pattern to new stimuli is only possible if participants have learned a broad generalisation about identical and non-identical pairs, as opposed to learning a narrow generalisation about each individual segment pair. These results support the hypothesis that a distinction between identical and non-identical segments is learned as such, with direct reference to identity, as predicted by the copying model.

In addition to extending the identity-based pattern to novel stimuli in the CRITICAL condition, a difference between old and new stimuli was also found for the CONTROL condition. The interaction between condition and old *vs*. new found that participants in the CONTROL condition gave far fewer alternating responses to stimuli with new consonant pairs than to those with old consonant pairs. This effect is likely a side-effect of the bias against alternation, as discussed for Experiment 1. While participants in the CONTROL conditions are exposed to 50% alternating and 50% non-alternating stimuli pairs in training, they only give alternating response around 33% of the time in testing. For stimuli with new consonant pairs, there is no evidence from training that those pairs should or could alternate, and so the bias against alternation appears even more strongly.

6 Conclusion

This paper has presented results from two artificial grammar learning experiments that provide support for the explicit representation of identity in phonological generalisations. The experiments show that a pattern based on the distinction between identical and non-identical consonant pairs is easier to learn than an arbitrary pattern, and that an identity-based pattern generalises to novel consonant combinations. Both of these results suggest that learners make a single generalisation over identical/non-identical segments, as opposed to tracking only the patterning of individual segmental or featural pairs.

The copying model is successful at modelling the experimental results and the typology of phonological patterns, in that it predicts both that the identity pattern should be learned better than an arbitrary pattern, and that it should extend to novel segment pairs. The success of the copying model rests on a particular representational assumption: in the input to the learner, identical segments have only a single feature matrix. In this way, the model is similar to autosegmental accounts of the identity effect (McCarthy 1986, 1988), which assume that the only representational possibility for identical segments (or features) is a single set of features linked to multiple prosodic positions. In the copying model, the unique representation of identical segments restricts the ability of constraints to apply to both identical and non-identical segment pairs, and in turn allows distinct constraints on identical and non-identical segments to be learned. Whether this particular assumption of the copying model is a realistic model of how identity-based patterns are learned is a substantial question for future research. The questions for a model is (i) whether the distinction between identical segments is in the representation of the learning data or in the generalisations that are learned over the data and (ii) whether identity in either representations or constraints is a bias that learners come with, or is learned from the data.

While the copying model assumes that the identity relation is in the input data, Berent *et al.* (2012) pursue the idea that it is the generalisations over the learning data that reference identity. They propose an expansion to the constraint-induction algorithm in the UCLA phonotactic learner to include constraints that refer to variables. In this model, there is nothing special about the representation of identical segments in the input data – they are represented as sequences of two feature matrices, just like any other pair of segments. Constraints, however, may penalise either simple sequences of feature matrices, e.g. *[+cg][+cg] penalises a sequence of [+constricted glottis] segments, regardless of other features of the segments, or variable-based relationships between feature matrices, e.g.

$*[+cg]_i[+cg]_i$ penalises a sequence of [+constricted glottis] segments only if they are identical. To learn an identity-effect pattern, this type of model would need to consider constraints with both variable-matching relationships and variable-mismatching relationships, i.e. the constraint necessary for the identity effect is $*[+cg]_i[+cg]_j$. It remains to be tested whether a model that induces variable-based constraints can make the same predictions as the copying model, but some speculations can be offered. It seems likely that such a model would, like the copying model, predict that an identity-based pattern is learned better than an arbitrary pattern, since the identity pattern can be stated with a single, general constraint $*[+cg]_i[+cg]_j$. What is less clear is whether a model with variable-based constraints can generalise an identity-effect pattern to novel identical segments. In the simulations in §3, it was found that the baseline model predicted an unattested identical ejective pair ([p'…p']) to be even worse than the unattested non-identical ejective pairs. This prediction was made because the unattested ejective pair could be penalised by constraints that ruled out non-identical pairs; the relevant high-weighted constraint was *[+cg, LAB][+cg]. This constraint would also be available in a model with variable-based constraints, and it may thus lead such a model to predict that the identity pattern would not extend to novel segments. It is left to future work to verify and explore the predictions of this type of model.

Whether the identity relation in either representations or constraints is induced from the data or part of a prior bias has implications for languages that don't exhibit identity-based patterns. In addition to languages with the identity effect, many languages have co-occurrence restrictions that apply equally to identical and non-identical segments, like Quechua, which disallows all combinations of identical or non-identical ejectives in a root. In the copying model, this restriction would necessarily be learned as two separate generalisations, one, *[+cg][IDENT], over identical pairs of ejectives and the other, *[+cg][+cg], over non-identical pairs. If there is no bias towards identity, and identity is only included in constraints or representations when explicitly supported by the data, then a language like Quechua could be learned with just a single constraint on all pairs of ejectives, *[+cg][+cg]. Ongoing experimental work by the author is investigating whether speakers of languages like Quechua show a latent distinction between identical and non-identical segments. If true, this would be a case of a cross-linguistic tendency showing up at the level of an individual speaker, as has been found in acceptability-rating tasks for unattested onsets in English (Scholes 1966, Treiman *et al.* 2000, Albright 2009), as well as in a range of production and perception tasks (Broselow & Finer 1991, Broselow *et al.* 1998, Moreton 2002, Hansen 2004, Davidson 2006, Berent *et al.* 2007).

A bias in favour of identity may be analytic, in Moreton's (2008) terms, in that the formal structure of an identity-based generalisation is preferred. The preference for more general constraints that favour the identity pattern over the arbitrary pattern, discussed in §3, is a kind of analytic bias, but it may also be that learners have a bias in favour of identity-based generalisations. This would predict that learners prefer generalisations that reference identity or other simple algebraic relations over otherwise equivalent generalisations that are stated over simple featural strings. Support for a bias in favour of identity is presented in Endress *et al.* (2007), who find that a tonal pattern based on identity is learned better than a systematic tonal pattern that does not contain an identity relation. Bias may also be substantive. Previous experiments in artificial grammar paradigms have found that phonetically grounded patterns are learned better than phonetically unnatural patterns (Wilson 2006, Carpenter 2010, Finley 2011), and it may be that a preference for identical segments is phonetically grounded, though this remains to be shown experimentally.

This paper has outlined two specific properties that a model of learning must have: it must allow an identity-based generalisation to be learned more strongly than an arbitrary generalisation, and it must allow an identity-based pattern to generalise to novel items. While it is clear that identity can be represented explicitly by learners, and that languages can exploit the identical/non-identical distinction in grammatical patterns, it is not yet clear precisely how identity is represented in the grammar, or how the learning of identity-based patterns should be modelled.

REFERENCES

Abramson, Arthur S. & Leigh Lisker (1970). Discriminability along the voicing continuum: cross-language tests. *Proceedings of the 6th International Congress of Phonetic Sciences*. Prague: Publishing House of the Czechoslovak Academy of Sciences. 569–573.

Albright, Adam (2009). Feature-based generalisation as a source of gradient acceptability. *Phonology* **26**. 9–41.

Aulie, H. Wilbur & Evelyn W Aulie (1978). *Diccionario ch'ol–español, español–ch'ol*. México, D. F.: Institutó Lingüístico de Verano.

Bates, Douglas M. & Martin Maechler (2010). Package 'lme4' (Version 0.999375-34): linear mixed-effects models using S4 classes. Available (April 2013) at http://cran.r-project.org/web/packages/lme4/lme4.pdf.

Becker, Michael & Jonathan Levine (2010). Experigen: an online experiment platform. Available (April 2013) at https://github.com/tlozoot/experigen.

Berent, Iris, Gary F. Marcus, Joseph Shimron & Adamantios I. Gafos (2002). The scope of linguistic generalizations: evidence from Hebrew word formation. *Cognition* **83**. 113–139.

Berent, Iris & Joseph Shimron (1997). The representation of Hebrew words: evidence from the Obligatory Contour Principle. *Cognition* **64**. 39–72.

Berent, Iris, Donca Steriade, Tracy Lennertz & Vered Vaknin (2007). What we know about what we have never heard: evidence from perceptual illusions. *Cognition* **104**. 591–630.

Berent, Iris, Colin Wilson, Gary F. Marcus & Douglas K. Bemis (2012). On the role of variables in phonology: remarks on Hayes and Wilson 2008. *LI* **43**. 97–119.

Berg, René van den & La Ode Sidu (1996). *Muna–English dictionary*. Leiden: KITLV Press.

Broselow, Ellen & Daniel Finer (1991). Parameter setting in second language phonology and syntax. *Second Language Research* **7**. 35–59.

Broselow, Ellen, Su-I Chen & Chilin Wang (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* **20**. 261–280.

Buckley, Eugene (1990). Edge-in association and OCP 'violations' in Tigrinya. *WCCFL* **9**. 75–90.

Carpenter, Angela C. (2010). A naturalness bias in learning stress. *Phonology* **27**. 345–392.

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

Coetzee, Andries W. (2009a). Grammar is both categorical and gradient. In Steve Parker (ed.) *Phonological argumentation : essays on evidence and motivation*. London: Equinox. 9–42.

Coetzee, Andries W. (2009b). Learning lexical indexation. *Phonology* **26**. 109–145.

Coetzee, Andries W. & Joe Pater (2008). Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *NLLT* **26**. 289–337.

Colavin, Rebecca S., Roger Levy & Sharon Rose (2010). Modeling OCP-Place in Amharic with the Maximum Entropy phonotactic learner. Ms, University of California, San Diego. Available (April 2013) at http://idiom.ucsd.edu/~rose/Modeling%20OCP.pdf. To appear in *CLS* **46**.

Davidson, Lisa (2006). Phonology, phonetics, or frequency: influences on the production of non-native sequences. *JPh* **34**. 104–137.

Endress, Ansgar D., Ghislaine Dehaene-Lambertz & Jacques Mehler (2007). Perceptual constraints and the learnability of simple grammars. *Cognition* **105**. 577–614.

Finley, Sara (2011). The privileged status of locality in consonant harmony. *Journal of Memory and Language* **65**. 74–83.

Frisch, Stefan A., Janet B. Pierrehumbert & Michael B. Broe (2004). Similarity avoidance and the OCP. *NLLT* **22**. 179–228.

Frisch, Stefan A. & Bushra Zawaydeh (2001). The psychological reality of OCP-place in Arabic. *Lg* **77**. 91–106.

Gafos, Adamantios I. (1998). Eliminating long-distance consonantal spreading. *NLLT* **16**. 223–278.

Gallagher, Gillian (2010a). *The perceptual basis of long-distance laryngeal restrictions*. PhD dissertation, MIT.

Gallagher, Gillian (2010b). Perceptual distinctness and long-distance laryngeal restrictions. *Phonology* **27**. 435–480.

Gallagher, Gillian & Jessica Coon (2009). Distinguishing total and partial identity: evidence from Chol. *NLLT* **27**. 545–582.

Goldsmith, John A. (1976). *Autosegmental phonology*. PhD dissertation, MIT.

Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenador, Anders Eriksson & Östen Dahl (eds.) *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm: Stockholm University. 111–120.

Graff, Peter & Florian Jäger (2009). Locality and feature specificity in OCP effects: evidence from Ayamar, Dutch and Javanese. *CLS* **45**. 127–141.

Halle, Morris (1962). Phonology in a generative grammar. *Word* **18**. 54–72.

Hansen, Jette G. (2004). Developmental sequences in the acquisition of English L2 syllable codas: a preliminary study. *Studies in Second Language Acquisition* **26**. 85–124.

Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *LI* **39**. 379–440.

Hudson, Grover (1995). Phonology of Ethiopian languages. In John A. Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell. 782–797.

Jaynes, E. T. (1983). *Papers on probability, statistics, and statistical physics*. Edited by R. D. Rosenkrantz. Dordrecht: Reidel.

Jensen, John T. (1977). *Yapese reference grammar*. Honolulu: University Press of Hawaii.

Klein, Dan & Christopher Manning (2003). Maxent models, conditional estimation, and optimization, without the magic. Tutorial presented at NAACL-03 and ACL-03.

Leben, William (1973). *Suprasegmental phonology*. PhD dissertation, MIT.

Lisker, Leigh & Arthur S. Abramson (1967). Some effects of context on voice onset time in English stops. *Language and Speech* **10**. 1–28.

Lucca, Manuel F. de (1987). *Diccionario practico aymara–castellano castellano–aymara*. La Paz: Editorial Los Amigos del Libro.

McCarthy, John J. (1986). OCP effects: gemination and antigemination. *LI* **17**. 207–263.

McCarthy, John J. (1988). Feature geometry and dependency: a review. *Phonetica* **45**. 84–108.

McCarthy, John J. & Joe Pater (eds.) (to appear). *Harmonic Grammar and harmonic serialism*. London: Equinox.

MacEachern, Margaret R. (1999). *Laryngeal cooccurrence restrictions*. New York: Garland.

Manning, Christopher D. & Hinrich Schütze (1999). *Foundations of statistical natural language processing*. Cambridge, Mass: MIT Press.

Marcus, Gary F., S. Vijayan, S. Bandi Rao & P. M. Vishton (1999). Rule learning by seven-month-old infants. *Science* **283**. 77–80.

Mester, R. Armin (1986). *Studies in tier structure*. PhD dissertation, University of Massachusetts, Amherst.

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

Moreton, Elliott (2008). Analytic bias and phonological typology. *Phonology* **25**. 83–127.

Moreton, Elliott (2012). Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language* **67**. 165–183.

Pater, Joe (2009). Weighted constraints in generative linguistics. *Cognitive Science* **33**. 999–1035.

Pater, Joe, Rajesh Bhatt & Chris Potts (2007). Linguistic optimization. Ms, University of Massachusetts, Amherst. Available (April 2013) at http://people.umass.edu/pater/pater-bhatt-potts-hg07.pdf.

Pater, Joe & Elliott Moreton (2012). Structure and substance in artificial-phonology learning. 2 parts. *Language and Linguistics Compass* **6**. 686–701, 702–718.

Pierrehumbert, Janet B. (1993). Dissimilarity in Arabic verbal roots. *NELS* **23**. 367–381.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.

Rose, Sharon (1997). *Theoretical issues in comparative Ethio-Semitic phonology and morphology*. PhD dissertation, McGill University.

Rose, Sharon & Lisa King (2007). Speech error elicitation and co-occurrence restrictions in two Ethiopian Semitic languages. *Language and Speech* **50**. 451–504.

Sagey, Elizabeth (1986). *The representation of features and relations in nonlinear phonology*. PhD dissertation, MIT.

Scholes, Robert J. (1966). *Phonotactic grammaticality*. The Hague: Mouton.

Straight, H. Stephen (1976) *The acquisition of Maya phonology: variation in Yucatec child language*. New York & London: Garland.

Thomas, Jacqueline M. C. (1963). *Le parler Ngbaka de Bokanga : phonologie, morphologie, syntaxe*. Paris : Paillart.
Treiman, Rebecca, Brett Kessler, Stephanie Knewasser, Ruth Tincoff & Margo Bowman (2000). English speakers' sensitivity to phonotactic patterns. In Michael B. Broe & Janet B. Pierrehumbert (eds.) *Papers in laboratory phonology V : acquisition and the lexicon*. Cambridge : Cambridge University Press. 269–282.
Warkentin, Viola M. & Ruth M. Brend (1974). Chol phonology. *Linguistics* **132**. 87–101.
Wilson, Colin (2006). Learning phonology with substantive bias : an experimental and computational study of velar palatalization. *Cognitive Science* **30**. 945–982.