Phonology **27** (2010) 345–392. © Cambridge University Press 2010
doi:10.1017/S0952675710000199

*A naturalness bias in learning stress**

Angela C. Carpenter
Wellesley College

Recent research on the acquisition of natural *vs*. unnatural phonological processes provides some support for the idea that learning a natural process is easier than learning an unnatural one (Wilson 2003, 2006, Pycha *et al*. 2003, Pater & Tessier 2005). This study extends those findings by comparing the acquisition of two stress patterns that are identical except in naturalness. Learners were native speakers of English, a language with variable stress, and French, a fixed stress language. Both English and French speakers learned the natural pattern significantly better than the unnatural. The artificial languages specifically neutralised the phonetic cues that might have given a perceptual advantage to the natural language. The findings suggest that a naturalness bias aids in the distinguishing and learning of a phonological pattern. To explain the results, I argue for an interaction between a general and a language-specific cognitive mechanism.

1 Introduction

A fundamental issue in phonological theory is whether speakers have knowledge of phonological universals, or whether generalisations about possible *vs*. impossible (or rare *vs*. common) phonological systems are simply the products of historical change. In the current literature, the position that speakers have knowledge of universals is typically associated with Optimality Theory (OT; Prince & Smolensky 1993), while the product of change view is associated with Evolutionary Phonology (Blevins 2004). Each of these works contains a long list of cited precedents for the universalist and emergentist viewpoints; see also Wilson (2006) and Moreton (2008) for intermediate positions.

* I would like to thank three anonymous reviewers and an associate editor of *Phonology* for their comments, which has considerably improved the content and presentation of this paper. I am also grateful to Lawrence Baldwin and Arnold Well for their statistical help. In addition I wish to thank Andries Coetzee, Paul de Lacy, John Kingston, John McCarthy, Tom Roeper and Shelley Velleman for their input and suggestions in the preparation of this work, and Lori Morris and UQAM for hosting me during data collection in Montreal. A special thanks to Joe Pater and Andrea Levitt for their input on several versions of this paper. This research was supported in part by grants from the Ford and Mellon Foundations and a post-doctoral fellowship supported by Wellesley College. Any errors are solely mine.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

A recently popular technique for probing knowledge of universals is artificial language learning. If speakers have knowledge of universals, then a pattern that satisfies them should be easier to learn than a comparable one that does not. This technique was pioneered by Schane *et al.* as a test of the universalist assumptions of Natural Phonology (Stampe 1979) and has re-emerged with the advent of OT. The evidence from these studies is mixed. Some researchers have found that learning a natural pattern is easier than learning an unnatural one (e.g. Schane *et al.* 1974, Healy & Levitt 1980, Pater & Tessier 2005, Peperkamp *et al.* 2006, Zhang & Lai 2010), while others show no difference in learning between natural and unnatural rules (e.g. Pycha *et al.* 2003, Peperkamp & Dupoux 2007). Of those that found a difference, some of the studies have incorporated features of naturalness that could be confounded by language learners' previous exposure to those natural patterns in their native language (e.g. Schane *et al.* 1974, Healy & Levitt 1980). Other experiments avoided the issue of learners being previously exposed to the target natural features, but the comparison was not strictly natural *vs.* unnatural; rather, they could be explained as being natural *vs.* less natural (e.g. Pycha *et al.* 2003, Wilson 2006).

Part of the reason that the evidence is mixed is that the terms 'natural' and 'unnatural' are used in a wide variety of ways. A natural rule is sometimes categorised as one that is formally simple (e.g. manipulates a single feature), as one that is phonetically or otherwise substantively grounded, or as one that is robustly attested cross-linguistically (see Pycha *et al.* 2003, Wilson 2003 and Pater & Tessier 2005 for discussion). In this study, I examine two stress rules that are formally equivalent, but differ along a substantive dimension, which distinguishes them in terms of being attested cross-linguistically. One rule stresses the first low vowel, else the first vowel, while the other stresses the first high vowel, else the first vowel. Low vowels tend to attract stress because of their sonority (Kenstowicz 1994). Stress rules that target low vowels are attested cross-linguistically; ones that target high vowels are not (de Lacy 2002). I find that both English and Quebec French subjects perform better in learning the natural pattern than the unnatural one. Since both patterns are to some degree learned, it is appropriate to refer to this as a NATURALNESS BIAS towards learning the stress pattern (Wilson 2006, Moreton 2008).

1.1 Second-language learnability

The study carried out in this paper involves artificial language learning by adults, which is a form of second-language (L2) learning. Two of the proposals made to account for first-language learnability within OT are the Gradual Learning Algorithm (Boersma 1997) and Constraint Demotion (Tesar & Smolensky 1998, 2000). While Constraint Demotion theory and the Gradual Learning Algorithm could possibly account for learning a second language, neither theory makes an explicit claim to account for L2 learning. Given the limitations of these constraint-based

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

models for second language learning, I am simply making the general assumption that the language-learning mechanism includes constraint reranking for both children and adults. Further, the mechanism is capable of interacting with other cognitive processes, as suggested by research showing that adults have some learning advantages over children in structured language learning settings (Olson & Samuels 1973, Snow & Hoefnagel-Höhle 1982, Hansen 1995). The exact nature of the cognitive mechanism is unclear, but it could take the form of a generalised mechanism that also works for language, as in statistical language learning (e.g. Saffran *et al*. 1996, Saffran & Thiessen 2003).

Thus it is probable that a language-learning task involves using a combination of learning mechanisms: some influenced by an innate mechanism specific to language and others influenced by general cognitive mechanisms. In comparing the learning of a natural *vs*. an unnatural stress pattern, I make the assumption that learners use both the general cognitive mechanism and a language-learning specific mechanism. The language-specific mechanism should produce a naturalness bias, i.e. a tendency to better learn the natural rule over the unnatural one. This is because the natural rule is specific to language, while the unnatural rule is not, so that a language-specific mechanism can be of help in learning the linguistic rule. If a language-specific mechanism is not available to help in second-language learning, then subjects should learn the natural and the unnatural languages equally well, simply by using general cognitive mechanisms. If language-specific cognition is available, then the learners of the unnatural language are at a disadvantage because the unnatural version goes against a natural linguistic principle, and innate knowledge of a linguistic principle will not help in that case. Note that I do not predict that the unnatural version is unlearnable, just that it will not be learned as well as the natural version. In summary, both languages are learnable, but a cognitive ability that is specific to language gives the natural language an advantage.

The rest of the paper is organised as follows: §2 lays out the formal theoretical background as well as the role of phonetics in the study, §3 details the vowel-height experiments conducted with American English speakers, showing that the natural stress rules were better learned than the unnatural ones, §4 discusses the results of the same study done with Quebec French speakers, thus replicating the results with learners from a different (fixed) stress system, §5 discusses an alternative explanation for learning the natural language and §6 concludes the paper.

2 Theoretical background

2.1 Natural stress systems and vowel height

As mentioned in the introduction, while the term 'natural' has a number of interpretations, the distinctions between natural and unnatural rules that I examine are clear-cut: the natural rule is cross-linguistically attested

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

and the unnatural rule is not attested in language. The natural rule I use in the experiments is grounded in the linguistic distinction of vowel height. By changing a feature of the natural principle to its opposite, an unnatural rule is created that produces a phonological pattern that is unattested in language. The important result of this rule manipulation is that two rules can be pitted against each other, both of which are formally similar and thus should be equally learnable. However, this similarity in formality does not take into account naturalness in language, thus setting up an experimental situation where learnability of a natural *vs*. an unnatural rule can be observed.

In OT, universal principles are instantiated as a set of violable constraints, common to all languages, which are ranked within each language in a way that accounts for its observed phonological patterns. Given a universal set of constraints, a natural language results from a particular ranking of those constraints. A basic assumption of OT is that each set of rankings represents a different language or dialect, producing a factorial typology of all languages. A factorial typology of these constraint rankings, that is, all licit permutations of all the constraints, should yield all the natural stress patterns in natural languages and none of the unnatural ones. 'Licit permutation' excludes rankings that go *against* universally fixed hierarchies. Fixed hierarchies are those in which the constraints cannot be freely permuted with respect to each other (Prince & Smolensky 1993, McCarthy 2002). Fixed hierarchies are based on natural scales that hold universally for all languages. For example, vowels are more sonorous than consonant stops, or first person is more prominent than third person. Translating natural scales into OT requires two or more constraints to be in a fixed, non-permutable, domination relation. For example, if constraints A and B are in a fixed hierarchy, where A ranks higher than B, resulting in a natural prominence relationship, the converse ranking, where B ranks higher than A, is not a possible permutation, since the resulting linguistic feature would not be found in natural language.

A factorial typology of constraints could produce stress patterns that are possible but unattested, but being unattested alone would not make a language unnatural. It is also possible that while known constraints together produce all the stress patterns in the world's languages, an as yet unknown constraint could also account for some undiscovered stress pattern. Clearly, this is also not what would be classified as an unnatural pattern. To be unnatural, the language would have to go against one or more of the following: theoretically possible constraints that are determined by universal linguistic principles of markedness and faithfulness, a fixed universal hierarchy or a stringency relationship based on markedness principles.[1]

[1] I thank an anonymous reviewer for pointing out the important distinction between an unattested language and a truly unnatural or impossible language.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Under Optimality Theory, then, a natural/unnatural pair of stress systems can be derived in one of the following ways: (a) the natural system conforms to a universally fixed hierarchy, while the unnatural system goes against such a hierarchy, or (b) the natural system results from the ranking of well-motivated constraints, while the unnatural system requires a constraint that has no basis in known linguistic principles of markedness or faithfulness, in other words, a non-linguistic constraint. The unnatural stress rule devised for these experiments requires stressing of a high over a low vowel, which goes against a fixed constraint ranking, the Peak Prominence hierarchy (Prince & Smolensky 1993).

2.2 Stress based on vowel height

It has been observed that stress seeks high sonority segments (Kenstowicz 1994, de Lacy 2002). Although all vowels have high sonority relative to consonants, within the vowel system itself some vowels are of higher sonority than others. This sonority hierarchy can be expressed as a scale where low vowels are more sonorous than mid vowels, which are in turn, more sonorous than high vowels and glides (see Clements 1990, who credits Sievers 1881 and Jespersen 1904 for early development of these principles). For some few languages, stress placement is, in part, determined by the sonority of the vowel. In Gujarati, for example, stress is attracted to a low vowel [a] rather than to mid and high peripheral vowels or schwa. The following rules are adapted from the descriptions of Cardona (1965) and de Lacy (2002).

(1) *Gujarati stress*

Default stress is on the penultimate syllable.

a. If syllables have the same vowel, or if the penultimate syllable contains a vowel other than [a], stress occurs on the penultimate syllable.

['bidi]	'a type of cigar'	['bolo]	'old'
['kedi]	'prisoner'	['bɛso]	'sit down'

b. However, stress will move to a non-penult with [a].

[de'khav]	'appear(s)'	[nuk'san]	'damage'
['taʤetər]	'recently'	[sine'ma]	'cinema'

c. Otherwise, stress a non-final, non-low vowel [ɛ ɔ e o i u].

[nə'muno]	'sample'	[mə'lisuN]	'we will meet'

Another language in which vowel height serves to attract stress is the Moksha dialect of Mordvin (Finno-Ugric, spoken in the Volga region of Central Russia). Moksha Mordvin requires stress on the left edge of the word, unless the leftmost syllable is a 'narrow' vowel [i u ə], in which case stress occurs on the leftmost 'broad' vowel [e o æ ɑ] (Cygankin & Davaev 1975, Kenstowicz 1994).

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

(2) *Moksha Mordvin stress*

Stress leftmost broad (non-high) vowel, else leftmost.

a. If all vowels are narrow, or all vowels are broad, stress the initial syllable.

['puvəndəms]	'to press'	['kizəfnʲəms]	'to press'
['ræmɑsɑk]	'you buy it'	['noldɑsɑk]	'you release it'

b. If the word contains both narrow and broad vowels, stress the first broad vowel.

['sɑʃəndɑt]	'you arrive'	[tuʃən'dɑt]	'you go away'

This type of pattern, where primary stress takes vowel height into account, is relatively rare in the world's languages. I use such a pattern as the basis for this artificial language experiment and test the strength of the ability to recognise a possible (but rare) phonological pattern against one that is impossible.

The Peak Prominence hierarchy (Prince & Smolensky 1993), a universal scale, plays a key role in accounting for the preference for high-sonority (low) vowels as syllable nuclei in languages such as Moksha Mordvin and Gujarati. As natural languages, they do not violate the universality of the scalar relationships within the Peak Prominence hierarchy. Other languages in which sonority plays some role in stress placement include Alyawarra (Yallop 1977), Armenian (Vaux 1998), the Pichis dialect of Asheninca (Payne 1990, Hayes 1995), Chukchee (Krause 1980, Kenstowicz 1994), Finnish (Hayes 1995), Komi (Itkonen 1955, Hayes 1995) and Nganasan (Helimski 1998, de Lacy 2002, 2004).[2] In support of sonority as a factor for attracting stress to low vowels over non-low vowels, de Lacy argues that 'the opposite "anti-Gujarati" situation never occurs: there is no language in which stress ignores [a] and seeks out a non-low vowel instead' (de Lacy 2002: 2). A language where stress is attached to a non-low vowel instead of a low vowel, all other factors being equal, is not only unattested, but unnatural.

2.3 The artificial natural and unnatural languages

The natural stress system I have adopted for this study follows the rule: stress leftmost low vowel, else stress leftmost vowel. In this language, a simpler version of Moksha Mordvin, stress emerges on the first syllable of words in which either (i) the first syllable contains a low vowel, such as ['pɑtikæ] and ['pætɑkɑ]; or, (ii) all the syllables contain high vowels, ['pitiku]. Stress is on the second or third syllable with a low vowel if the preceding syllable contains a high vowel, such as [pu'tɑkæ] and [pitu'pɑ].

[2] Although note that in English, under certain circumstances, stress will move away from a syllable with a sonorant consonant coda (Kiparsky 1979, Kager 1989).

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

The unnatural version of this stress pattern follows the rule: stress leftmost high vowel, else stress leftmost vowel. It is the mirror image of the natural version, in that stress occurs on the first syllable of words in which either (i) the first syllable contains a high vowel, such as ['putæki] and ['pituku], or (ii) all the syllables contain low vowels, ['pætɑkɑ]. Stress occurs on the second or third syllables with high vowels when preceding syllables contain a low vowel, such as [pɑ'tuki] and [pætɑ'pi].

2.4 Analysis of the high/low vowel languages

The purpose of the detailed analysis in this section is to show how Optimality Theory predicts natural stress patterns and rules out unnatural ones. The analysis of both the natural language and the unnatural one is centred on the principles of sonority and alignment. Following Prince & Smolensky (1993), I adopt the Peak Prominence hierarchy, where high-sonority segments are preferred over low-sonority segments as syllable nuclei. The hierarchy is based on the sonority scale, where low vowels are most sonorous and voiceless obstruents least sonorous.

Prince & Smolensky (1993) propose the following fixed constraint ranking, which captures the role of the sonority hierarchy in choosing optimal syllable nuclei. Here and below, T represents voiceless stops, D voiced stops, N nasals, L liquids, R rhotics, I high vowels, ə mid vowels and A low vowels.[3]

(3) $*P/T \gg *P/D \gg \ldots \gg *P/I \gg *P/A$

Smith (2002) uses the Peak Prominence hierarchy to build a universally fixed constraint ranking of stressed syllable nuclei. The resulting hierarchy, which I will refer to as the Stressed Peak hierarchy, formalises the preference for stressed syllable peaks that are maximally sonorous.

(4) $[*P_{K/T}]/'\sigma \gg [*P_{K/D}]/'\sigma \gg [*P_{K/N}]/'\sigma \gg [*P_{K/L}]/'\sigma \gg [*P_{K/R}]/'\sigma \gg [*P_{K/I}]/'\sigma \gg [*P_{K/ə}]/'\sigma \gg [*P_{K/A}]/'\sigma$

The crucial ranking used in the analysis that follows is: $[*P_{K/I}]/'\sigma \gg [*P_{K/A}]/'\sigma$. For expositional clarity, these constraints are referred to here as *HighStress and *LowStress, with the ranking *HighStress ≫ *LowStress.

[3] De Lacy (2004, personal communication) argues for a set of stringency constraints, based on markedness scales, that does not require a fixed, non-permutable ranking. Under this approach, the natural language would emerge under the ranking (slightly modified from de Lacy's) $*H_{D_{Ft}}/i\text{-}u \gg$ AlignHead-L, $*H_{D_{Ft}}/i\text{-}u,a$. The unnatural version could never emerge, because there is no constraint that would rule out stress on low vowels that would not also rule out stress on high vowels.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

(5) a. *HighStress
Stress must not occur on high vowels.
b. *LowStress
Stress must not occur on low vowels.

The tableau in (6) demonstrates how this hierarchy picks a stressed low-vowel syllable over a stressed high-vowel one. H indicates a high-vowel syllable and L a low-vowel syllable.

(6)

/HLL/	*HighStress	*LowStress
a. ˈHLL	*!	
☞ b. HˈLL		*

To produce a pattern where stress occurs on the *leftmost* syllable with a low vowel as the nucleus, there is an interaction between the Stressed Peak hierarchy and an AlignHead-L constraint.

(7) Align-L(Head,PWd)
The left edge of the head of the Prosodic Word must coincide with the left edge of the Prosodic Word.

Alignment constraints are gradiently violable, in that each instantiation of non-alignment within a given domain is assigned a violation mark (although cf. McCarthy 2003). So, using AlignHead-L as an example, if the head of the Prosodic Word (the stressed syllable) is two syllables away from the left edge then AlignHead-L is violated twice and receives two violation marks. In effect, AlignHead-L requires all words to have stress on the initial syllable. Ranking AlignHead-L below *HighStress results in stress occurring on the leftmost non-high vowel.

(8) *High and low vowels: stress leftmost low*

/pitaka/	*HighStress	AlignHd-L
☞ a. piˈtaka		*
b. pitaˈka		**!
c. ˈpitaka	*!	

As seen in tableaux (6) and (8), *HighStress dominates both *LowStress and AlignHead-L. The tableaux in (9) and (10) demonstrate how the ranking of AlignHead-L below *HighStress prevents stress from occurring on a high vowel at the left edge. This same ranking also correctly allows default stress to occur at the left edge.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

(9) *All high vowels: stress leftmost*

/pitiku/	*HighStress	AlignHd-L	*LowStress
☞ a. ˈpitiku	*		
b. piˈtiku	*	*!	
c. pitiˈku	*	*!*	

(10) *High and low vowels: stress leftmost low*

/pitaka/	*HighStress	AlignHd-L	*LowStress
a. ˈpitaka	*!		
☞ b. piˈtaka		*	*
c. pitaˈka		**!	*

The above tableaux demonstrate the effects of the alignment constraint within the Stressed Peak hierarchy. This effect of controlling how far away from the edge of a domain a segment or process can go has been well documented (e.g. McCarthy & Prince 1993, Kenstowicz 1994).

The unnatural stress system is just the opposite of the natural one: stress the leftmost *high* vowel, else stress the leftmost. The only way to obtain this output is to switch the ranking of the Stressed Peak hierarchy to unnaturally penalise low vowel nuclei, as in *LowStress ⩾ *HighStress. Such a ranking favours the high vowels [i] and [u] over low [ɑ] and [æ] as stressed syllable nuclei.

(11) *Unnatural stress system*

/patuki/	*LowStress	*HighStress
a. ˈpatuki	*!	
☞ b. pɑˈtuki		*

Ranking *LowStress above *HighStress produces a winning candidate where stress occurs on a high vowel instead of a low vowel. As in the natural version of this language, AlignHead-L restrains the stressed syllable from straying too far to the right. There is no ranking argument that shows a dominance relationship between AlignHead-L and *HighStress, therefore they are equally ranked. The constraint ranking then that produces the unnatural language is *LowStress ⩾ AlignHead-L, *HighStress.

I claim that this ranking is unnatural, as it goes against the Stressed Peak hierarchy. There is no supporting evidence from natural languages that disfavours higher-sonority syllable nuclei over lower-sonority ones, when all other factors are held equal (Prince & Smolensky 1993, de Lacy 2002, Smith 2002). In summary, the natural stress pattern is allowed under a constraint ranking that respects the Stressed Peak hierarchy, but the unnatural language illicitly switches the proper order of the hierarchy.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

2.5 The role of phonetics

Phonological patterns often have an underlying phonetic basis. For example, the Gujarati pattern of stress on low vowels can be explained phonetically: low vowels are intrinsically longer in duration than high vowels, and length is a correlate of stress. Therefore stress is attracted to the longer vowel. This phonetic fact is supported by acoustic studies that have shown that the vowels in stressed syllables are of longer duration than that of unstressed syllables (e.g. Fry 1955, Jassem *et al*. 1968, Fant *et al*. 1991). In addition to duration, other correlates of stress are intensity, pitch and perceptual energy. In creating the stimulus materials, I adjusted duration and intensity for the vowels to compensate for the extra length intrinsic to the low vowels and the greater intensity of the high vowels. (See the methods section for details of these adjustments.)

Perceptual energy is also a correlate of stressed syllables, primarily in distinguishing between heavy and light syllables (Gordon 2002). Gordon gives evidence from Chickasaw, Khalkha and Finnish that stressed heavy syllables have more energy than unstressed ones. Energy, a factor of syllable duration and intensity, is defined here as in (12).

(12) $\int x^2(t)\, dt$

(x(t) is the amplitude of the sound at time t.) Perceptual energy then adds to the saliency of stressed syllables. In addition to neutralising the phonetic cues that favour perception of stress on low vowels, there was greater energy on high-vowel syllables than on low vowel ones.[4] Thus the perceptibility of stressed high-vowel syllables was slightly increased.

This disparity in energy actually overcompensates for the advantage that low-vowel syllables intrinsically have, further negating the phonetic differences between stressed high- and low-vowel syllables. Therefore, if learners are better able to identify and learn the natural low vowel stress pattern better than the unnatural high vowel pattern, then it is reasonable to conclude that something more than phonetics is influencing that ability. Thus we can begin to tease apart explanations of natural phonological patterns which rely on an influence of phonetics in learning from those which rely on the use of abstract representations of stress not involving phonetics.

2.6 Effects of the native language on learning the target languages

Since the experiments involve learning a non-native stress pattern, it is important to take a brief look at the effect of the native language in the learning process. As discussed in the L2 acquisition literature,

[4] §3.2.2 details the manipulations made to the original recorded syllables which increased the energy of high-vowel syllables.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

L2 learners exhibit effects of interference from the native language (Mairs 1989, Archibald 1995, Pater 1997; see also Broselow *et al*. 1998). By design, both the target stress patterns are similar to each other, but very different from English. However, it could be argued that native English speakers have some advantage in learning, because their native language has variable stress, as do the target languages. In the following sections I discuss the relevant facts about English stress that demonstrates what English speakers bring to the task of learning the target stress rules.

2.6.1 *English stress*. English stress has been much studied in the generative tradition (e.g. Chomsky & Halle 1968, Hayes 1982, Halle & Vergnaud 1987, Kager 1989, Burzio 1994, McCarthy & Prince 1995, Pater 2000, to mention just a few). I will not attempt a new analysis of English stress here, but will merely state the basic facts about the stress system as they pertain to this experiment with English-speaking participants.

Primary stress in English is not uniform, and can occur on a number of different syllables, depending in part on the syntactic category of the word. The stress pattern of nouns is of particular interest, since the words in the experiment are all nouns. In English nouns, stress can either occur on a heavy penultimate syllable, such as *ho*ˈ*rizon* and *u*ˈ*tensil* or, in the case where the penultimate syllables are light, on the antepenult, as in *A*ˈ*merica* and ˈ*cinema*. While main stress is usually on one of the final three syllables of the word, in some exceptional cases where secondary stress comes later in the word, primary stress can occur even earlier, through stress retraction, as in ˈ*catamaˌran* and ˈ*alaˌbaster*. In addition to pre-antepenultimate, antepenultimate and penultimate stress, English also has final stress on nouns with a long final vowel, as in *briˈgade*. Thus English permits stress on a range of syllables and stress placement is highly variable, though principled.

3 Experiment 1: natural *vs*. unnatural with English speakers

This experiment and the ones following use nonsense words in an artificial language learning paradigm, following the precedent of earlier research, beginning with Esper (1925), which established the utility of using artificial languages to understand universal properties and structures of language. Esper (1925) and other studies have demonstrated that phonological patterns can be learned very quickly in a laboratory setting (e.g. Schane *et al*. 1974, Dell *et al*. 2000, Onishi *et al*. 2002, Pater & Tessier 2005, Wilson 2006). After participants demonstrate that they have grasped the stress pattern on artificial words they hear, they are then tested on new nonsense words not previously heard. This methodology allows us to see if subjects have truly internalised the target stress pattern to the extent

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

that they can apply the acquired pattern to novel test words. Using novel words as test items neutralises the objection that stress placement is simply memorised, since subjects would have never heard those words before.

3.1 Research questions and prediction

The research question behind this experiment is this: is learning a natural phonological stress pattern easier than learning an unnatural one? The prediction is that learners of the natural stress pattern will do better than learners of the unnatural pattern, regardless of their native language. The indicator of success will be learners' scores on novel words after being trained in the natural and unnatural patterns respectively.

3.2 Method

3.2.1 *Design*. Subjects were randomly assigned to learn either the natural or the unnatural language. They were trained, using feedback, to correctly choose a correctly stressed over an incorrectly stressed version of the same nonsense word in a two-alternative forced-choice task (2AFC). A 2AFC task presents participants with two choices of the stimuli, a correct and an incorrect version; the participants then have to choose which version is correct. Feedback was provided to improve learnability, much as one would experience in classroom L2 learning. Participants were then tested twice on novel words, once in the middle of the training procedure and once at the end of the training using the same 2AFC procedure. The dependent variable was the proportion of correct choices made on the novel word tests.

Training/familiarisation. The training (or familiarisation) required the learner to simply listen to the nonsense words and look at a picture corresponding to each word. Pilot testing had previously shown that subjects learned best when the words heard were accompanied by a visual referent (cf. Hayes-Harb *et al*. 2010). The referents were photographs of real-life items, such as a bicycle, a mask or a tyre, designed to provide a meaning for each word and also to establish that the nonsense words were nouns. Knowing that the words were nouns would help native English speakers to pay attention to the variability in the stress pattern, something they are accustomed to hearing in English nouns. The artificial word used to represent each picture bore no similarity to the English word for the object pictured.

After each training block, subjects were tested on the now familiar training words in a 2AFC task. After hearing the two word choices, participants had to follow the instruction on the monitor: *Press the RED key* (*1*) *if the 1st word sounds right, or press the BLUE key* (*2*) *if the 2nd word sounds right*. The correct and incorrect versions were balanced as to order of presentation (correct then incorrect, or incorrect then correct) as well as to choice of which incorrectly stressed syllable was presented. For example, if participants heard ['bɑdɑkɑ – bɑ'dɑkɑ] (correct stress on the

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

first syllable, then incorrect stress on the second syllable) during Test 1 after Training 1, then they heard [bɑdɑ'kɑ – 'bɑdɑkɑ] (incorrect stress on the third syllable, then correct stress on the first syllable) when that training word pair was next tested.

Novel word testing. Immediately following the testing of their knowledge of the first 18 training words, and again after being exposed to all 27 training words, participants were tested on novel words, designed to see how well they had learned the stress pattern. The novel words repeated the stress pattern they were trained on, so participants demonstrated their understanding of the underlying pattern by scoring well on the novel words. The first novel word test consisted of 18 different word pairs. After choosing, subjects were given feedback with the correct answer. The second novel test was composed of 48 more novel words. No feedback was given during the final novel word test.

Participants faced a slightly more difficult task with the novel word tests than with the training words tests. The reasons are: (a) they had had no training on these words before, and (b) learners had to make an adjustment to being tested without the benefit of having an accompanying image. Thus we can conclude that being tested on the novel words was more difficult, because of the lack of training and a visual referent.

3.2.2 *Stimuli*. The stimulus set consisted of three- and four-syllable nonsense words made up of 32 different syllables: eight different consonants (six stops and two fricatives), combined with four different vowels. The vowel inventory used here was composed of four vowels: the high vowels [i] as in *beat* and [u] as in *boot*; and the low vowels [æ] as in *bat* and [ɑ] as in *pot*. The consonant inventory was composed of the voiced and voiceless stops [b d g p t k] and the fricatives [s z].

To create the stimulus materials, the target syllables were embedded in a carrier sentence, where they received uniform stress. The carrier sentence was *Mike __ Karinna Jones*, with the target syllable inserted in the blank space. The sentences were recorded by a male trained phonetician, a native speaker of American English. The recorded stimuli were then digitised and each syllable saved separately. All syllables were scaled to fit a temporal envelope with a silent onset of 5 ms, and a 5 ms trailing edge; the maximum value in the file was scaled up so that its maximum amplitude was close to the maximum possible amplitude.

Two sets of changes were then made to the syllables, the first to neutralise the natural distinction in length and intensity between high and low vowels, and the second to increase the perceptible difference between stressed and unstressed syllables.

Some stressed syllables are inherently longer than others. For example, stressed [ɑ] and [æ] tend to be longer than [i] and [u]. This length difference (the longest syllables, those with [ɑ], were about 27% longer on average than those with [u]) could bias subjects towards preferring stress more readily on the longer vowels than on the shorter ones. Since the longer vowels are all low vowels, subjects learning the natural rule would

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

have a phonetic reason to prefer stress on the correct syllables. This phonetic advantage was removed by equalising vowel durations, while retaining the natural variation between syllables with different onsets. To equalise the vowels the length of all the vowels in each syllable by syllable group, e.g. [bɑ], [dɑ], [gɑ], [pɑ] and [kɑ], was measured first, then an average vowel length of the [ɑ] syllables was obtained. Average vowel lengths were obtained for all syllable groups, [ɑ], [æ], [i] and [u]. A grand average (average of the averages) was then obtained, which became the new target average for all the syllable groups. Calculating the difference between the grand average and the syllable group's average gave the percentage of increase or decrease needed on each syllable within a group to make all the overall syllable lengths equal. So, for example, all the [æ] syllables were decreased by 6·90 %, to obtain an average syllable length of 208·43 ms. In contrast, all the [i] syllables were increased by 8·94 %, to obtain the same average syllable length. These syllable manipulations were done in Praat (Boersma & Weenink 2006), using the duration manipulation and increasing the relative duration of the vowel only by using a multiplier. The syllables were then resynthesised with the Pitch Synchronous Overlap Add (PSOLA) algorithm.

Average intensities were measured over the entire syllable using Praat's query function. The natural intensities of high-vowel syllables were an average of 2 dB higher than that of low-vowel ones. In general, stressed high vowels are usually of lower intensity than stressed low vowels (Ladefoged 2003: 92). However, the original stressed high-vowel syllables, as recorded for these experiments, were of higher intensity than the stressed low-vowel syllables.

Calculations for the intensity adjustments needed were made as follows: average intensity was computed of all syllables that began with the same consonant. A grand average of these averages produced a percentage by which the high-vowel syllables, [Ci] and [Cu], were to be reduced. This reduction produced an overall equalisation of intensity of all the syllables. The actual intensity manipulation was made by creating an intensity tier in Praat and multiplying over the entire syllable. In the final stimuli, intensities of the low-vowel syllables are nearly identical to the high-vowel syllables. For example, all [bi] syllables (unstressed, initial, medial and final stressed) were reduced by about 2 dB, whereas the corresponding [bu] syllables were reduced by about 2·7 dB. The overall difference in intensities between the set of low- and high-vowel syllables was less than 1 %.

The adjustments described above produced a set of base syllables that were equalised in duration and intensity. To increase the perceptibility of stress on the stressed syllables, further adjustments were made to amplitude, duration and pitch. Previous research has shown that a change in pitch on the vowel results in an increased judgment of stress on the syllable (Fry 1955). So, to maximise stress perceptibility, a pitch contour was added to each stressed syllable. In summary, to create stressed syllables, the following changes were made to the base syllables: (a) duration of

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

the vowel was increased by 20%; (b) pitch was increased by 20% and a pitch contour was added based on the syllable position: syllables in the initial stress position of the word received a falling contour, syllables in the medial stress position received a rising then falling contour and syllables in the final stress position received a rising contour.

For unstressed syllables, intensity of the base syllables was reduced by 6 dB. These manipulations produced perceptually salient stressed syllables for both high- and low-vowel syllables. Manipulations to the high-vowel syllables also resulted in these syllables having more energy than the low-vowel ones. Pilot testing showed that with these manipulations subjects could reliably (97% of the time) pick out the stressed syllable in a three- or four-syllable nonsense word.

Words for both the natural and unnatural languages were concatenated using the 'Combine Sounds' feature in Praat from the same set of stressed and unstressed CV syllables. 27 training words and 66 novel words were constructed. Of the 27 training words, 12 had three syllables and 15 four syllables. The possible word types for each group of three- and four-syllable words was 2^n, i.e. eight for three-syllable words and 16 for four-syllable words. For three-syllable words the possible word types were: HHH, HHL, HLL, HLH, LLL, LLH, LHL and LHH (where H = high vowel and L = low vowel). The distribution of stressed syllables in both the training and testing portions was made in a way that produced the default pattern of initial stress, while exposing the learner to equal numbers of non-initially stressed words. Therefore, of the three-syllable words, 50% had first-syllable stress, 25% second-syllable stress and 25% third-syllable stress. Within the four-syllable words, half had first syllable stress, and stress was divided evenly among the remaining three syllables (approximately 17% each). These proportions, 50% of the words having initial stress and with stress evenly divided among the remaining syllables by word type, were maintained in the 66 novel test words, so that the test ratio imitated the learning ratio.

To make both languages as similar as possible, words for both natural and unnatural versions used the same consonant sequences, but with different vowels, depending on whether they were in the natural or unnatural group. The vowels were interchanged according to the following pattern: back vowels [ɑ] for the natural and [u] the unnatural; and front vowels [æ] for the natural and [i] for the unnatural. Thus consonant sequences and the frontness/backness of vowels and were matched across languages. To illustrate, the first four training words in the natural language were ['bɑdɑkɑ], with the consonant sequence [bdk], ['gikusi] ([gks]), [zu'pætɑ] ([zpt]) and [tigu'sɑ] ([tgs]). The corresponding training words in the unnatural language were ['buduku] ([bdk]), ['gækɑsæ] ([gks]), [zɑ'pitu] ([zpt]) and [tægɑ'su] ([tgs]). Novel words were matched in the same way (see Appendix A for a list of the words used).

3.2.3 *Participants.* All participants were adult native speakers of English with no known hearing problems, from Wellesley College and the

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

University of Massachusetts, Amherst. While students all had exposure to other languages through classroom language instruction, English was their native language. The call for participants requested monolingual English speakers. No one reported being bilingual. Subjects were either paid for participation or received class credit. All participants from Wellesley College were students in psychology classes who were required to participate in an experiment. The students from University of Massachusetts, Amherst were recruited through their classrooms and through the linguistics department's online experiment sign-up procedure.

Of the 44 participants, four were excluded: one did not meet the pre-training criterion and three did not meet the criterion of getting 70% correct on the test of all 27 training words. There remained 21 subjects in the natural language group and 19 in the unnatural one. Subjects had no previous experience with the Moksha dialect of Mordvin, on which the stress principle of the natural version of this artificial language was based.

3.2.4 *Procedure*. Participants took part in the experiments either individually or in small groups of two or four on individual computers. They were told that the purpose of the experiment was to understand how English speakers learned a stress pattern that was different from their native stress system. All were given an instruction form that explained the nature of the experiment, what stress was and how the experiment would proceed. They were also informed that the words they were to learn were nonsense words that included the following vowels: [æ] as in *hat*, [ɑ] as in *hot*, [i] as in *beet* and [u] as in *boot*. After being given an opportunity to ask questions, each sat at a computer station and donned headphones. They were told to follow the instructions on the computer screen in order to proceed through the experiment.

Participants were presented with a pre-training block to make sure they could hear the distinction between stressed and unstressed syllables. The pre-training block presented words in an AXB task, where participants heard nonsense words in groups of three. Words A and B were stressed on a different syllable and X matched either A or B as to stress. The task was to indicate whether either the first two words or the second two words had stress on the same syllable. In the first six trials the words presented in the AXB pattern differed only in the location of stress, for example ['biguti – 'biguti – bi'guti]. In the remaining 20 trials, subjects had to pick out the matching stress pattern from words that were completely different segmentally, such as [gi'zumi – pitɑ'bɑ – sɑti'bi]. Those who successfully identified the correct stressed syllable at least 70% of the time over all 26 trials were included in the main experiment.

Following the pre-training, all participants were randomly assigned to either the natural or the unnatural version of the main experiment. Clear instructions as to how to proceed were presented on the computer

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

monitor. The main experiment consisted of two parts: (a) a training or familiarisation component in which participants were exposed to the words along with pictures and then tested on just those familiarisation words, and (b) testing on novel words. All tests during the training and the novel word testing were presented as a two-alternative forced choice task. After hearing the two choices, written instructions on the computer monitor told them to indicate which word had the correct stressed syllable, word 1 or word 2.

The order of the experiment was as follows: during the training/familiarisation component, participants heard the correct version of each word, while at the same time they saw a photographic image corresponding to that word. Each sound/image combination lasted for one second, and the same combination was randomly repeated four times in each training block. Training/familiarisation blocks alternated between presenting four three-syllable words and then five four-syllable words. After hearing a block of words, participants were given a 2AFC test on that particular block of words. During these tests, participants received feedback informing them of the correct choice, regardless of what their answer was. The feedback consisted of seeing the phrase *The correct answer is*:, and then hearing the correct version of the word. Following the first two training/testing blocks, they listened to a review of these first nine words, repeated twice, and were tested on those words. The process was repeated with nine new words, where they were exposed to a new group of four three-syllable words and a group of five four-syllable words, and then tested on those words. They then reviewed the first 18 words by listening to them repeated twice and were tested on those 18 words. At this point it was expected that they would have begun to learn the first 18 words, having had a total of 12 exposures to the first 18 words. They were then tested on 18 novel words that they had never heard before. This first novel word test included feedback. Following this test of novel words they were once again trained and tested on nine new words, then heard a review of all 27 words, and were tested on all the training/familiarisation words. The final test was made up of 48 novel words. The format of the experiment is given in Table I.

The experiment was conducted using SuperLab Pro, which recorded subjects' responses to each test item. Participants' responses were then saved in an Excel file. Correct and incorrect answers were converted to numerical values of 1 and 0 respectively, and proportion correct was calculated as the dependent variable. Proportions based on binary choices can lead to a non-normal distribution, thus undermining statistical analyses that assume a normal distribution. Thus, once the data were transferred into SPSS 17.0 for analysis, an arcsine transformation was computed of the proportions ($y = \sin^{-1}(\sqrt{(x)})$, where x is the original proportion) (cf. Wilson 2003). The statistical analyses reported below were computed on these transformed data. However, for ease of understanding, I show both the actual and the transformed percentages in the tables.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Pre-training	Subjects complete AXB test: 26 triplet groups
Training block 1	Subjects hear 4 three-syllable training words, randomly repeated 4 times, each word represented by a unique photographic image.
Testing block 1	Subjects tested (2AFC) on the 4 three-syllable words just heard. They get feedback with the correct answer after responding to each pair presented.
Training block 2	Subjects hear 5 four-syllable training words, randomly repeated 4 times, each word represented by a unique photographic image.
Testing block 2	Subjects tested on the 5 four-syllable training words just heard, with feedback.
Review	Subjects hear the first 9 training words, presented randomly and repeated once, i.e. each word heard twice.
Testing block 3	Subjects tested on the first 9 training words, with feedback.
Training/ testing blocks 4 and 5	Each training block presents new words and subjects are tested on just the words presented in that training block. Same procedure as blocks 1 and 2 above.
Review	Subjects hear the second 9 training words, presented randomly and repeated once, i.e. each word heard twice.
Testing block 6	Subjects tested on the second 9 training words, with feedback.
Review	Subjects hear the first 18 training words, presented randomly and repeated once.
Testing block 7	Subjects tested on the first 18 training words, with feedback.
First novel words test	Subjects presented with 18 novel three- and four-syllable test words. They get feedback as to the correct answers.
Training/ testing blocks 8 and 9	Same procedure as training/testing blocks 1, 2, 4 and 5 above. Training and testing block 8 presents and tests a new set of three-syllable words and block 9 trains and tests a new set of four-syllable words.
Review	Subjects hear the third set of 9 training words, presented randomly and repeated once.
Testing block 10	Subjects tested on the third set of 9 training words, with feedback.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Review	Subjects hear all 27 training words, presented randomly and repeated once.
Testing block 11	Subjects tested on the 27 training words they have learned, with feedback. This test measures how well they have learned the training words.
Final novel words test	Subjects presented with 48 novel test words, with no feedback.

Table I
Experiment format.

3.3 English results

3.3.1 *Learning the training/familiarisation words.* Participants needed to learn the training words as a way of becoming familiarised with the target stress pattern. How well they learned those words was measured by their score on the test of all 27 training words (testing block 11 in Table 1). On the test of 27 words, the natural language group scored a mean of 90% correct and the unnatural group scored a mean of 89% correct. One-sample *t*-tests on the arcsine transformed means were conducted to evaluate whether each mean was significantly different from the expected chance level, 50%. Both the natural and unnatural groups scores were significantly different from chance (natural group: transformed mean = 1·25, SD = 32·46, $t(20) = 10{\cdot}655$, $p < 0{\cdot}001$; unnatural group: transformed mean = 1·18, SD = 30·80, $t(18) = 9{\cdot}670$, $p < 0{\cdot}001$). A one-way ANOVA by subject was conducted to determine whether there was a significant difference in score between the natural and the unnatural groups. The independent variable was the experimental language, natural or unnatural, and the dependent variable was the arcsine transformed means of the test of all 27 training words for participants. The ANOVA showed that there was no significant difference between the scores of the natural and unnatural language learners. Thus both groups were well familiarised with their respective stress patterns.

3.3.2 *Generalising the stress pattern: performance on the novel word tests.* The real test of how well participants learned the underlying stress pattern of the language they were exposed to was measured by their performance on the novel words. The score was computed as the percentage correct of all the novel words. The natural group scored 70·40% correct and the unnatural 61·97% correct. Both groups' scores on the novel words were significantly better than chance, as shown by the results of a one sample *t*-test on the arcsine transformed proportions for each group chance (natural group: transformed mean = 80·30, SD = 20·49, $t(20) = 6{\cdot}77$, $p < 0{\cdot}001$; unnatural group: transformed mean = 68·33, SD = 18·89, $t(18) = 4{\cdot}23$, $p = 0{\cdot}001$). Table II presents the raw scores and scores of both the natural and unnatural groups.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

English speakers (novel words)	Raw score (% correct)	SD	Arcsine transformed score (% correct)	SD
natural group	70·40 $p < 0·001$	12·11	80·30 $p < 0·001$	20·4 9
unnatural group	61·97 $p < 0·05$	12·90	68·33 $p < 0·05$	18·89

Table II

Correct novel words (percentage) and Standard Deviation: English speakers.

	Stress position					
	Initial (%)	SD	Medial (%)	SD	Final (%)	SD
natural group	69·6 5	11·93	72·86	11·57	68·6 6	9·74
unnatural group	68·42	8·77	56·84	11·01	55·6 4	11·81

Table III

Means and Standard Deviation for novel words correct by stress position: English speakers.

As stated in §3.1, I predicted that the natural group would perform better on the novel words than the unnatural group. To test this prediction, an independent samples *t*-test was conducted on the arcsine transformed scores to evaluate whether the means of the natural group were significantly different from those of the unnatural group. The dependent variable was the percentage correct on the novel words and the independent variable was the language, natural or unnatural. The natural group's score of 70·40% correct (arcsine M = 80·30%, SD = 20·49%) was significantly higher than the unnatural group's 61·97% (arcsine M = 68·33%, SD = 18·89%), $t(38) = 1·915$, $p < 0·05$ (one-tailed).

A between-items 2×3 ANOVA was conducted on the novel words to evaluate the effects of language and three stress positions on subjects' mean scores. The two language factors were natural and unnatural, and the three stress positions were initial, medial and final. Medial position includes the second syllable in three-syllable words and the second and third syllables in four-syllable words. The untransformed means and standard deviations for scores on the novel words are presented in Table III. The results for the ANOVA indicated significant main effects for language, $F(1, 126) = 26·312$, $p < 0·001$, partial $\eta^2 = 0·17$; stress position, $F(2, 126) = 4·477$, $p < 0·05$, partial $\eta^2 = 0·07$; and a significant interaction between stress position and language, $F(2, 126) = 6·703$, $p = 0·002$, partial $\eta^2 = 0·10$. The significant effect for language confirmed the between-subject results in the independent samples *t*-test that, given a choice of correctly stressed and incorrectly stressed novel words, subjects made the correct choice in the natural language significantly more often than in the unnatural language.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Scores in the natural language did not differ significantly according to the stress position: initial, medial or final. That was not the case with the other group. In the unnatural language, subjects scored best on initial stress, but less well on medial and final stress, as illustrated in the graph in Fig. 1. Because the interaction between stress position and language was significant, follow-up tests were conducted to evaluate the pairwise differences among the means of the unnatural language with respect to stress position. To control for Type I error and to reduce the possibility of a family-wise error in making a post hoc analysis of the three pairwise comparisons, alpha was reduced to 0·0125. There were significant differences within the unnatural group between initial and medial stress, $F(1, 126) = 14{\cdot}634$, $p < 0{\cdot}001$; and between initial and final stress, $F(1, 126) = 13{\cdot}362$, $p < 0{\cdot}001$. There was no significant difference between medial and final stress. This result indicates that learners of the unnatural language were significantly better at identifying stress on the initial syllable than on medial and final syllables.

3.4 Discussion: English speakers

As predicted, the natural language group, with an average score of 70% correct on all the novel words, learned the stress pattern of their language better than the unnatural language group, which scored 62% correct. As a reminder, the natural language group learned a rule where stress occurred on low vowels, while the unnatural language group learned a rule where stress occurred on high vowels. Since a preference for stressed high vowels instead of low vowels, all things being equal, is unattested in language, then this result lends support to the idea that learners of the low-vowel language had a naturalness bias in their favour.

A look at the percentage of correct choices grouped by stressed syllable in Table III shows that English speakers in the natural group made correct choices at about the same level for initial, medial and final stress. This result indicates that this group did better at learning both the stress rule (stress leftmost low vowel) and the default (else leftmost vowel). On the other hand, the group that learned the unnatural language were fairly good at correctly choosing initial stress, but performed significantly worse on medial and final stress. Given that initial stress was the default, with half of all the words in the training or familiarisation phase having initial stress, we can conclude that participants' scores in the unnatural language group reflected the statistical frequency of the stimulus words. Thus it could be said that the unnatural group was better at learning the default than at learning the main stress rule.

Perhaps these data shed some light on the difference between use of a general cognitive mechanism for learning and a language-specific mechanism. Both groups had general cognitive mechanisms available to aid in learning the pattern of the target language. However, due to the experimental manipulation, only the natural language group could have utilised a language-specific mechanism to tacitly gain a better sense

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

of the target language's main stress rule. The general mechanism would allow them to take note of the frequency of initially stressed syllables and thus become more attuned to words with correct stress on the first syllable. However, exposed to data that goes against a natural language rule, it would have been more difficult for them to figure out that rule. I will discuss the implications of this distinction in the concluding section.

To summarise, Experiment 1 shows that English speakers were able to learn the natural rule, stress leftmost low vowel, else leftmost vowel, better than the unnatural one, stress leftmost high vowel, else leftmost. This result suggests that learners were influenced by a naturalness bias in learning stress. The question now arises, is this result robust across languages? Further, some might object by saying that there is some slight evidence that English speakers may already have a bias towards low vowels attracting stress. To answer these questions the experiment was run with native speakers of Quebec French. Quebec French was chosen for a number of reasons, which are detailed below.

4 Experiment 2: Quebec French speakers

Although by most accounts, English stress is not dictated by vowel height, Rice (1996) suggests that there is some slight evidence that stress is attracted to low vowels in English. Rice argues that irregular penultimate stress on certain English nouns can be analysed as an effect of stress occurring on low vowels. In general, light penultimate syllables do not bear primary stress. However, words such as *ba*ˈ*nana* and *go*ˈ*rilla* prove to be exceptions to this generalisation. His reasoning is this: high vowels cannot attract a second mora without changing their phonemic identity. For example, a monomoraic [ɪ], if augmented with another mora, would then correspond to its tense counterpart [i]. Low vowels, however, with an additional mora, do not become 'categorically different' (Rice 1996: 166). Thus, by Rice's account, the low vowels of *ba*ˈ*nana* and *Ala*ˈ*bama* attract penultimate stress by virtue of being bimoraic.

Rice's argument raises the possibility that English speakers have tacit knowledge that low vowels are special (also see Hammond 1999). Thus, to counteract this argument, I ran the experiment with French speakers, for which no such claim has been made.

Another argument suggesting that low vowels could be perceived as being special for English speakers is this: since low vowels are inherently longer in duration than high vowels and duration is associated with the perception of stress, English speakers could have internalised a preference for stress to occur on low vowels. So even though the vowel lengths were equalised across both natural and unnatural languages, it is possible that English speakers could still use the tacit knowledge of vowel length to better aid them to learn the natural rule over the unnatural one. Using speakers of Quebec French serves to counteract this possible English

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

advantage, since French has a fixed stress system wherein stress is consistently on the last non-schwa syllable. The following sections provide more details about stress in French.

Finally, Quebec French was also chosen for reasons of phonetic similarity to American English. It has a similar vowel inventory to American English, with a high/low opposition in two places, like the target languages. Quebec French has the front and back low vowels [æ] and [ɑ] and the front and back high vowels [i] and [u] (Tranel 1987, Dechaine 1990). Further, the consonants used in the experiment languages are also present in Quebec French.

4.1 French stress

The standard description for the placement of main stress in both European and Quebec French is that stress is fixed on the final (non-schwa) syllable in multisyllabic words (Schane 1968, Dell 1980, Charette 1991).[5] If the final syllable ends in a schwa, then primary stress occurs on the penultimate syllable. However, dialectal differences between European and Quebec French sometimes affect stress assignment. In Quebec French (QF), main stress is sometimes shifted to the penultimate syllable, perhaps due to vowel lengthening (Dechaine 1990) or syllable weight (Paradis & Deshaies 1990, Armstrong 1999).[6] While these prosodic analyses raise the possibility that stress in QF could have some variability, the preponderance of research still shows that QF has main stress on the right edge of the word. Thus QF is a good choice for a language with a different stress pattern than English. Furthermore, the experimental languages favour stress on the left edge of the word, which is the opposite of stress attraction in Quebec French.

An even more significant characteristic of French speakers that affects this study is the phenomenon of so-called stress 'deafness'.

[5] It has been argued that French has no word stress *per se*, but instead has stress at the end of the phonological phrase (e.g. Delattre 1939, Walker 2001). However, when words are said in isolation, the prosodic pattern of final stress at the end of the phonological phrase holds at the word level.

[6] Paradis & Deshaies, in a stress-perception experiment with speakers from the Chicoutimi-Jonquière area of Quebec, calculated that 'the general probability of a syllable being perceived as stressed is ·28 and ·24', although it could fall to 0·04 (1990: 142). So when there was stress on a syllable, subjects could only accurately pick that syllable as being stressed at best 28% of the time, thus confirming that speakers of Quebec French have difficulty in consistently identifying stress. Within those low probabilities, when vowel height was factored in as a stress attractor, stress was perceived more often on low and mid-high vowels than on high and mid-low ones. High vowels and schwa were the 'least favorable environments' for stress attraction (Paradis & Deshaies 1990: 142). This finding actually supports the universal principle of stress not being attracted to high vowels.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

4.2 French stress 'deafness'

Recent research has delved into how perception of prosodic differences between languages influences the acquisition of a second language, and specifically, the acquisition of a native-like accent. The inability of native speakers of French to reliably perceive stress, a phenomenon sometimes referred to as stress 'deafness', has been a relatively robust finding in recent perception literature (Dupoux *et al*. 1997, Peperkamp *et al*. 1999, Peperkamp & Dupoux 2002).

Dupoux *et al*. (1997) first reported on a series of experiments showing that, in comparison to Spanish speakers, French speakers made significantly more errors and were significantly slower in identifying the stressed syllable in nonsense words. This limited capacity to perceive a phonological property of a foreign language has been termed phonological 'deafness' (Peperkamp *et al*. 1999). Peperkamp *et al*. (1999), using a different methodological paradigm, had similar findings, where French subjects made twice as many errors as Spanish subjects.

Because of these findings, native French speakers needed to be accommodated in some way in order to ensure that they were able to learn the new stress patterns of the experimental languages. This accommodation will be discussed in the methodology section (§4.5.4).

4.3 Predictions

The overall prediction for these experiments is that the natural language rule will be more easily acquired than the unnatural rule. To acquire the natural language, French speakers will have to rerank the Stressed Peak hierarchy and AlignHead-L with respect to AlignHead-R, so that stress will occur on the leftmost syllable with a low vowel. The ranking hierarchy that produces this outcome is *HighStress ⩾ AlignHead-L, *LowStress. To produce the unnatural version, the opposite ranking must be produced, namely *LowStress ⩾ AlignHead-L, *HighStress. This ranking results in stress occurring on the leftmost high vowel, else leftmost. As previously discussed, this ranking goes against the natural hierarchy, where stress occurs on more sonorous nuclei, and should be more difficult for the learner to formulate. In contrast, the natural language's constraint ranking gives the learner a linguistic advantage in acquisition. The prediction then is that the French speakers will learn the natural version of the vowel height language better than the unnatural version. Specifically, the natural group will score higher on the novel words than the unnatural group.

Given the cross-linguistic nature of the study, a comparison will be made between English attainment and French attainment of both the natural and unnatural languages. In order to successfully learn the natural language, both French and English speakers have to activate the Stressed Peak hierarchy, as vowel height does not play a primary part in stress assignment in either language. In OT terms, the AlignHead-L

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

constraint, which is crucial in the target languages, is low-ranked in both French and English. Those two considerations appear to put the two languages on an even playing field in terms of ability to learn the natural version. However, the English speakers' variable stress system is expressed by a larger set of active constraints than that which produces the French stress system. In itself, this idea does not predict whether English learners would have an advantage in learning the target natural language. However, the English system, like the target languages, does allow stress in various positions in the word, while French is much more limited. The intuition is that this difference could give English speakers an advantage in acquiring the variable natural language. The prediction, therefore, is that French speakers will not learn the natural and unnatural languages as well as English speakers.

4.4 Experiment 2A: pilot testing for stress 'deafness'

Although the experimental methodology and the stimuli were the same for French participants as for English, a different pre-training procedure had to be devised for French speakers to accommodate European French speakers' diminished ability to reliably perceive stress (Dupoux *et al.* 1997, Peperkamp & Dupoux 2002), as discussed above.

First, to determine whether speakers of Quebec French (and not just European French) had difficulty in reliably choosing a stressed syllable, a pilot experiment was run with a separate group of French speakers. Participants were presented with an AXB discrimination task, in which subjects, after hearing a sequence of three stimulus words, had to say whether the second stimulus (X) had stress on the same syllable as the first (A) or the third (B). Subjects were presented with 25 different three- and four-syllable words made up of CV syllables. These syllables were created as described in §3.1.2, i.e. pitch and duration of stressed syllables were increased by 20%.

4.4.1 *Procedure*. The experiment was constructed with the use of SuperLab Pro software and presented on a laptop computer. A total of 25 trials were presented in two different blocks. The first block presented six trials where all three words were segmentally the same, but the stress was on a different syllable, as in ['dipɑgɑ – di'pɑgɑ – di'pɑgɑ]. The following block consisted of 19 trials, where all three words were different and subjects had to choose the pair of words that had stress on the same syllable, such as [tibɑkɑ'gu – tuti'duzu – bizu'pæsi]. Participants listened to the words through headphones, while simultaneously viewing the following instruction on the monitor: *Indicate which two words are similar. Press the RED key if the 1st two words are similar. Press the BLUE key if the last two words are similar*. Subjects could take as much time as they needed to respond. Once they responded, the next trial was presented. There was an inter-stimulus interval of 500 ms before each set of words were presented. All blocks were randomised for each subject. Each block would

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

begin only after subjects pressed the spacebar, thus they had an opportunity to rest between blocks if they so chose. Responses were recorded and stored in a data file.

Seven native French speakers were tested in this pilot. The results showed that, on average, they correctly identified the stressed syllable on the first attempt only 51% of the time. Table IV shows the individual results. This average of 51% shows that French speakers were not able to reliably identify the stressed syllable when those syllables only had a 20% increase in duration. In comparison, native English speakers ($n = 11$) correctly picked out the stressed syllable 93% of the time. Since the ability to correctly pick out the stressed syllable in nonsense words is crucial to participants' ability to learn the natural and unnatural artificial languages, it was necessary to make an adjustment that would allow French speakers to reliably identify stress in each word.

Subject	Correct
1	52%
2	44%
3	55%
4	43%
5	56%
6	31%
7	73%
average	51%

Table IV
Correct stressed syllables (percentage) at 120% duration for French speakers.

Previous research has shown that duration is a strong correlate of stress for French speakers (Fant *et al.* 1991). When syllable duration was increased from 120% of the equalised lengths to 180%, French speakers' accuracy improved greatly. However, it was important that both English and French participants heard the same stimuli in the main experiment. Thus, the pre-training portion described in the next section was designed to train French speakers to move from needing the 180% duration to being able to reliably hear and identify stress at the 120% duration level. After being trained to identify stressed syllables at the 120% level, they could then participate in the main experiment.

4.5 Experiment 2B: natural and unnatural language experiment

4.5.1 *Method.* Experiment 2B is a repetition of Experiment 1 using French speakers, but with a different pre-training portion than the one used in Experiment 1 with English speakers. As mentioned above, the purpose of the pre-training was to make sure participants could reliably

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

identify stressed syllables. After completing the pre-training, participants went on to complete the exact same main experiment as English speakers did, with the exact same stimuli. The only difference was that the instructions and directions on the computer monitor were in French, not English.

4.5.2 *Participants.* Participants were college-age native speakers of Quebec French, who were students at the University of Quebec at Montreal.[7] None reported being bilingual or having any hearing problems. Students were recruited in linguistics classes and through posters in public spaces on campus, and were paid for their participation. A total of 45 French speakers participated in the experiment. Of these, three did not meet the criterion of being able to correctly identify stress on at least 70% of the pre-training words. Two others were omitted because they failed to meet the 70% criterion on the 27 familiarisation words. This left a total of 40 participants, 20 in the natural group and 20 in the unnatural group.

4.5.3 *Procedure for pre-training.* The goal of this part of the experiment was to train French speakers to perceive stress at the 120% duration level. To that end, three sets of stimuli were created: 60 words where the stressed syllable had 180% duration (long duration), 65 words with 150% duration (medium duration) on the stressed syllable and finally 68 words with 120% duration (short duration). The stimulus words used here were unique to the pre-training, and were not repeated in the main experiment. A total of 193 trials, one word per trial, were presented in three different blocks, with the long duration syllables in the first block, and the medium and short duration syllables in the following two blocks. Participants listened to the words through headphones while simultaneously viewing the instruction on the monitor: *Press 1, 2, 3 or 4 to indicate the stressed syllable.* Subjects could take as much time as they needed to respond; however, if they made an incorrect choice by hitting the wrong number corresponding to the correct stressed syllable, then the word would be repeated until they made the correct choice. There was an inter-stimulus interval of 500 ms before each word was presented. All blocks were randomised for each subject. Each block would begin only after subjects pressed the spacebar; thus they had an opportunity to rest between blocks if they so chose. Responses were recorded and stored in a data file. Averages were taken of the total number of responses in each block. Thus errors made on individual words reduced the overall percentage of correct responses for each group of words.

[7] Although Montreal is a very cosmopolitan city, in which many languages are spoken, French is the dominant language and the first language of 80% of the population. See Statistics Canada, 'Population by mother tongue, by province and territory (2006 census)'; available (September 2010) at http://www40.statcan.ca/l01/cst01/demo11b-eng.htm.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

The overall results of the pre-training was that participants' ability to correctly identify the stressed syllable increased, as seen in the results in Table V.

Stressed syllable length		Correct
long duration	(180%)	88%
medium duration	(150%)	89%
short duration	(120%)	87%

Table V

Correct stressed syllables (percentage) during pre-training for French speakers.

Clearly, gradually familiarising subjects with shorter and shorter durations helped them to better identify stressed syllables at the 120% level. Recall that French speakers who did not undergo this stress perception training scored a mere 51% correct on the short duration words.

French speakers who scored 70% or more correct on the shorter duration words were included in the main experiment. This criterion was the same as for native English speakers. After the pre-training, participants immediately proceeded to the main part of the experiment, which was exactly as the one presented to English speakers, the exception being that instructions and feedback were presented in French, not English.

4.5.4 *Design and stimuli.* The design and stimulus materials were exactly the same as used in the English experiment. The vowels used were the same as for English speakers, but Quebec French speakers identify them as [æ] as in *à* 'at', [ɑ] as in *pâte* 'dough', [i] as in *lit* 'bed' and [u] as in *sous* 'under'.

4.6 French results

4.6.1 *Learning the training/familiarisation words: French speakers.* As with the English speakers, it was important that French speakers learned the training words in order to do well on the novel words. Overall, French participants successfully learned the stress patterns of the target language, as seen by their scores on the test of all 27 training words. The natural group scored an average of 93% correct on the test of all 27 training words and the unnatural group scored 86%. A one-sample *t*-test on the arcsine transformed data for each group showed that participants learned their respective target languages at a level better than chance (natural group: arcsine $M = 131{\cdot}11\%$, $SE = 6{\cdot}33\%$, $t(19) = 12{\cdot}820$, $p < 0{\cdot}001$; unnatural group: arcsine $M = 106{\cdot}39\%$, $SE = 4{\cdot}62\%$, $t(19) = 12{\cdot}216$, $p < 0{\cdot}001$). While both groups learned the target languages well, the natural group with 93% correct of the training words did significantly better than the unnatural group, which scored 86% correct, according to results of an independent samples *t*-test ($t(38) = 3{\cdot}156$, $p < 0{\cdot}05$).

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

French speakers (novel words)	Raw score (% correct)	SD	Arcsine transformed score (% correct)	SD
natural group	59·62 $p < 0·001$	8·92	64·39 $p < 0·001$	11·70
unnatural group	53·71 $p < 0·05$	6·6 9	56·88 $p < 0·05$	7·85

Table VI
Correct novel words (percentage) and Standard Deviation: French speakers.

	Stress position					
	Initial (%)	SD	Medial (%)	SD	Final (%)	SD
natural group	64·33	17·14	57·14	16·06	56·50	18·4 6
unnatural group	54·83	13·83	54·64	10·69	55·00	12·92

Table VII
Means and Standard Deviation for novel words correct by stress position: French speakers.

4.6.2 *Generalising the stress pattern to the novel words: French speakers.* French subjects' ability to generalise the stress pattern was measured by their average score on the novel words. Participants averaged 59·62 % correct of the novel words on the natural version and 53·71 % correct on the unnatural. Both groups' score on the novel words were significantly higher than chance, as can be seen from the results of a one sample *t*-test on the arcsine transformed proportions for each group (natural group: transformed mean = 64·39, SD = 11·70, $t(19) = 5·50$, $p < 0·001$; unnatural group: transformed mean = 56·88, SD = 7·85, $t(19) = 3·92$, $p < 0·05$). Table VI presents raw scores and transformed scores of both the natural and unnatural groups.

An independent samples *t*-test conducted on the arcsine transformed proportions showed a significant difference between the natural and unnatural scores. The dependent variable was the percentage correct on the novel words and the independent variable was the language, natural and unnatural. The natural group's score of 59·62 % correct (arcsine M = 64·39 %, SD = 11·70 %) was significantly higher than the unnatural group's 53·71 % (arcsine M = 56·88 %, SD = 7·85 %), $t(38) = 2·83$, $p < 0·05$.

A between-items 2 × 3 ANOVA was conducted on the novel words to evaluate the effects of language (natural and unnatural) and three stress positions (initial, medial and final) on subjects' mean scores (arcsine transformed). The untransformed means and standard deviations for scores on the novel words are presented in Table VII. The results for the ANOVA indicated a significant main effect for language, $F(1, 126) = 4·613$, $p < 0·05$, partial $\eta^2 = 0·04$, confirming the results of the independent

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

	English		French	
Target language	Training words	Novel words	Training words	Novel words
natural	90·40	70·40	93·07	59·62
unnatural	88·5 1	61·97	85·63	53·71

Table VIII

Correct training and novel words (percentage) for English and French speakers.

sample *t*-tests that the natural language scored significantly higher than the unnatural language. There was no significant effect for stress position, nor an interaction between language and stress position. These results indicate that, unlike English speakers, French speakers' scores were not significantly affected by the position of the stressed syllable.

4.7 Discussion: French speakers

Quebec French speakers' results on the novel words show that they were better able to learn the natural stress pattern than the unnatural one. Since vowel height is not a factor in French stress assignment, French speakers could not have been influenced by a stress rule from their native language. This result adds weight to the idea of a naturalness bias that aids learners in the acquisition of an unfamiliar stress pattern.

The French speakers scored 64%, 57% and 57% correct on initial, medial and final stress positions respectively. Although these scores were not significantly different from each other, it is interesting to note that among the natural language group, French speakers performed notably better on initial stress than on other stress positions. This group's score for initial syllable shows that they had moved away from the influence of the native language stress pattern, a result that supports previous findings for Quebec French speakers (Pater 1992). The unnatural rule group's scores by stress position were essentially flat, showing no preference for stress position. While French and English speakers learned the natural rule better than the unnatural, there were distinct differences between the groups, which I discuss in the following section.

4.8 Similarities and differences of attainment for English and French speakers

By the criterion that learners should score at least 70% correct on the test of all 27 training words, both English and French speakers learned the training/familiarisation words of the natural and unnatural languages well. English speakers scored an average of 90% and 89% on the natural and unnatural languages respectively on the test of all 27 training words, while French speakers averaged 93% and 86% (see Table VIII). French

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

speakers learned the natural language significantly better than the unnatural on the test of all 27 training words ($t(38) = 3{\cdot}156$, $p < 0{\cdot}05$); there was no significant difference between the English and French scores on the training words.

In successfully generalising the stress pattern, as evidenced by their performance on the novel words, English speakers did significantly better than French speakers. An ANOVA was conducted on the arcsine transformed data to measure whether the differences in scores between English and French speakers for the novel words was significant. The independent variables were the experiment language (natural or unnatural) and the native language (English or French). The dependent variable was the average score of all the novel words. The ANOVA showed main effects for the experiment language, $F(1, 76) = 7{\cdot}35$, $p < 0{\cdot}05$, and the native language $F(1, 76) = 15{\cdot}34$, $p < 0{\cdot}001$, but no significant interaction between the target language and the native language, $F(1, 76) = 0{\cdot}33$, $p = 0{\cdot}569$. The native language main effect shows that English speakers learned the experiment languages better than French speakers did. The main effect for experiment language showed that the natural version was learned better than the unnatural. This latter result demonstrates that the natural language carried an advantage that outweighed native language differences.

Given that both English and French speakers did better on the natural language, were there specifics of the natural language that were consistently learned better than others for all participants? To explore this question, a between-item analysis was carried out of the differences between the novel words. The novel words incorporated several factors in their design: stress occurred in any of three positions, initial, medial or final, and words had two lengths, three syllables or four syllables. A repeated-measures ANOVA was conducted on the 66 novel words with within-subjects factors being stress position (initial, medial and final) and word length (three-syllable or four-syllable), between-subject factors were language (English, French) and condition (natural, unnatural) and the dependent variable was the arcsine transformed scores of the novel words. Mauchly's test indicated that the assumption of sphericity was not violated; I report the results of the multivariate tests. There was no significant difference for stress position or word length. However, there was a near-significant effect for the interaction of stress position, language and condition, $V = 0{\cdot}074$, $F(2, 75) = 2{\cdot}977$, $p = 0{\cdot}057$, multivariate $\eta^2 = 0{\cdot}074$. The means and standard deviations are given in Table IX and illustrated in Fig. 1.

In addition, the repeated measures design confirmed the significant difference between English and French, $F(1, 76) = 14{\cdot}917$, $p < 0{\cdot}001$, and between natural and unnatural languages, $F(1, 76) = 8{\cdot}376$, $p < 0{\cdot}01$.

4.9 General discussion

French participants learned the natural language better than the unnatural one, replicating the results of the English speakers. The overall result of

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Language	Condition	Stress position	Raw means %	SD	Transformed means %	SD
English	natural	initial	70·52	14·54	81·95	26·16
		medial	74·15	17·27	89·52	31·85
		final	67·69	26·40	83·46	42·84
	unnatural	initial	68·77	17·15	78·89	25·36
		medial	59·02	17·79	66·60	28·74
		final	56·77	19·99	63·02	27·02
French	natural	initial	64·33	17·14	71·93	22·62
		medial	57·14	16·06	62·36	20·83
		final	56·50	18·46	63·20	23·67
	unnatural	initial	54·83	13·83	58·91	16·71
		medial	54·64	10·69	58·37	13·22
		final	55·00	12·92	59·18	16·67

Table IX
Overall percentage of novel words correct by stress position.

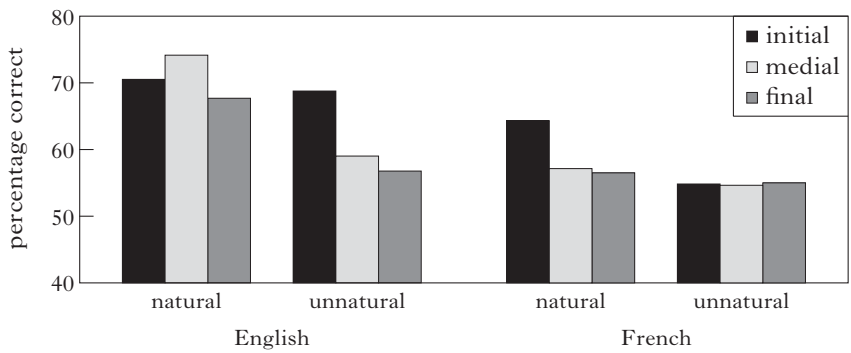


Figure 1
Novel words correct by target language, stress position and native language.

the comparison between English and French scores on the two languages supports the prediction that English learners would learn the natural language better than French learners: although the crucial Stressed Peak hierarchy is low-ranked in both English and French, English uses a larger set of active constraints to produce variable stress, which would give English subjects an advantage in acquiring another variable stress system.

The second prediction, that English participants would learn the unnatural language better than French learners, was also supported by the

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

data, although the size of the difference was larger than expected. While French speakers learned both target languages better than chance, their best performance (on the natural language) lagged behind the English speakers' worst performance (the unnatural language). Why was this? French performance could have been affected by several disadvantages. One is their inexperience with stress, which would have affected learning of both the natural and unnatural languages. Participants had to be trained to accurately identify stress, and they passed the necessary criterion indicating that they could do so. In addition, they learned the 27 training/familiarisation words (93% correct for the natural language and 86% correct for the unnatural one). While their performance more than adequately passed the criterion, the significant difference between the 93% and 86% indicates that the unnatural language was not learned as well as the natural. This could indicate an effect of an bias even in the learning of the unnaturally stressed training words.

Learning the training words required an active effort of memorising, with not much conscious effort going towards the underlying stress pattern. An indication of this difference comes from participants' own remarks about the process. A follow-up questionnaire administered to participants asked: *How would you describe your thought process in trying to figure out the pattern of the language?* And also: *How would you express the stress pattern of the language? In other words, how would you describe which syllable was stressed?* Most participants said that they learned the training words by memorisation. They also stated that they had no idea of the nature of the underlying stress pattern. The few that hazarded a guess were completely wrong. One participant's response to the latter two questions above was: *Je n'ai pas vraiment cherché de règle, je ne me concentrais que sur les repères visuels* ('I didn't really look for rules since I was concentrating so hard on the pictures'). I think that this statement reflects the nature of the learning task for most participants.

The overall diminished performance on the novel words could have resulted from the extra cognitive effort needed to pay attention to the task of hearing the stress *and* recalling the underlying stress pattern at the same time. Dupoux *et al*. found that in a simplified AX task, French subjects have 'very accurate access to the acoustic correlates of accent contrasts' (1997: 417). However, in a task that called for more cognitive capacity and use of working memory, their ability to accurately identify stress decreased. Their results suggest that in the experiments described here the task of internalising the stress pattern put a large cognitive load on French participants, which reduced their overall accuracy on the novel words. I attribute the difference between the natural and unnatural versions to a naturalness bias effect.

In addition to the cognitive load explanation, I speculate that the significant difference between French learning and English learning could be a result of another factor involving the nature of fixed stress. It could be that the constraint ranking that produces fixed stress is set fairly early in the language acquisition of French thus influencing how French speakers

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

store lexical representations of words. In contrast, stress in English is influenced by such a variety of factors that children do not acquire the full stress pattern until as late as six or seven years of age (Tom Roeper, personal communication). This difference between English and French speakers could translate to a cognitive difference in manipulating stress placement. English speakers continue reranking constraints to converge on their native stress pattern for a longer period of time during early learning. This increased and prolonged usage could serve to make the cognitive mechanism for learning stress more facile in native English speakers.

In the natural language group, English participants were most accurate at correctly identifying initial and medial stress and French participants at identifying initial stress, as demonstrated in Fig. 1. Everyone did relatively poorly on final stress. Although this result was not significant, there appears to be a trend towards stress early in the word. Could this trend be a result of their native language backgrounds, i.e. that English has a tendency towards initial and medial stress, while French has final stress? If native language were the main influencer for learning the natural language, then we would have to explain why there are opposite patterns: English learners tended to follow their native pattern while French learners tended to do the *opposite* of their native pattern. A more cohesive explanation lies in the nature of the stimuli presented to learners and the interaction between general cognitive learning and language-specific learning. Since the default stress was on the initial syllable, participants received more exposure to initially stressed words. This statistical fact could have been acquired through general cognitive learning principles. However, English learners' accuracy on medially stressed words demonstrates the emergence of their understanding of the main stress rule. I credit this learning to a language-specific mechanism. The story is a little less clear with French learners with respect to the interaction between general and language-specific cognition, although French learners were clearly going beyond their native stress pattern of final stress, since they scored 64% correct on words with initial stress and only 57% correct on final stress.

5 Learning through subvocalisation: an explanation for the results?

Participants were provided with several means of facilitating the learning of the training/familiarisation words and thus the underlying stress pattern.[8] These means were repetition, association of a picture with each word and feedback during the testing blocks. In addition to these overt efforts to help learning, it is possible, indeed likely, that participants also

[8] I thank an anonymous reviewer for pointing out the possible consequences of this learning mechanism.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

repeated the words to themselves as they heard them, a process termed subvocalisation, the study of which is a subfield of auditory imagery. Although this process of subvocalisation is most often reported in the literature with respect to reading (e.g. Baddeley *et al*. 1981, Abramson & Goldinger 1997), and the stimuli presented to participants in this experiment did not involve any reading, it is possible that participants used some form of subvocal repetition, especially when faced with the novel nonsense word task (see Hubbard 2010 for a thorough review of auditory imagery; also Thorn & Gathercole 2001, Aleman & van 't Wout 2004). A question which arises is whether participants changed the length in the low-vowel syllables when repeating the words to themselves. Doing so in the natural language would simply reinforce the pattern they were being taught. However, doing so in the unnatural version could possibly create competition between the target pattern of stress on high vowels and participants' underlying knowledge of length. If such were the case, then the decreased learning of the unnatural pattern could be due to the extra cognitive burden of these two competing factors.

Although there might have been possible effects of subvocalisation, it should be noted that the stressed high vowels were designed to have a high perceptual saliency. Equalising the duration of the vowels, along with increasing the overall duration, amplitude and pitch by 20%, resulted in the high-vowel syllables having more energy than the low-vowel ones. The mean energy, measured in Pa^2/sec, was 0·0199 for the low vowels and 0·0266 for the high vowels. An independent samples *t*-test showed that this difference was significant, $t(30) = -5{\cdot}176$, $p < 0{\cdot}001$. Since energy is a measure of salience, the conclusion is that the stressed high-vowel syllables should have been perceptually more salient than the stressed low-vowel ones. I believe that this saliency would serve to counteract any effect of subvocally increasing the duration of the low vowels. However, it is not entirely clear if the increased saliency was enough to create a just noticeable difference to the listeners (Watson 1973).

While English stress is not determined by a vowel-height distinction, the language has a durational difference between low vowels and high vowels. Since native English speakers in general have been exposed to this difference. it is conceivable that the native English speakers in this experiment could have added length to the low vowels when subvocalising. Given the nature of the stimuli and the experiment itself, I think this may not have happened, because the vowels in the 'unstressed' syllables were originally recorded in a stress position, and thus sound like full vowels, as they were not reduced. To add even more length to these vowels in repeating the words would have created words that were even more artificial sounding than they already were.

Speakers of Quebec French also exhibit a vowel-length difference in speech between high vowels and low vowels, where low vowels are longer than high vowels (Martin 2002). It is therefore possible that this group of French speakers could have added length to the vowels if they subvocalised. However, previous research has shown that stress-induced

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

segmental lengthening is 'of low perceptual salience' (Fant *et al.* 1991: 355) to French speakers. The preponderance of evidence points to speakers of Quebec French not being able to perceive stress very well (see note 6 and the results of the pilot study in §4.4). So while the low vowels may be longer in their speech, if this durational difference is not associated with stress, then they are probably not associating the length with stress, even if they put that length back during subvocalisation.

A follow-up questionnaire administered immediately after the experiment asked participants to explain how they learned the words. No one indicated that they said the words to themselves during any part of the experiment. Nevertheless, if participants did, in fact, subvocalise during the experiment and changed the length of low vowels, then that could be an alternate explanation for the experiment results, particularly for the lower scores of the unnatural language group. Specifically, by putting the duration of the low vowels back into the words as they subvocalised, participants would have created an environment where the stimulus presented with stress on high vowels would have been in competition with their own productions of stress on low vowels, thus creating confusion and decreasing the learning effectiveness.

To test this alternative a follow-up experiment could have a built-in design where participants have to repeat random numbers out loud in between trials, as a means of subverting subvocalising the experiment words. Another alternative, and one which is a next logical step to these experiments, is to have participants produce the words and then measure their relative vowel durations.

6 Conclusion

The empirical results from these experiments show that, cross-linguistically, a phonological pattern based on the natural linguistic principle of stress occurring on low vowels is better learned than a formally similar, but unnatural, principle of stress being attracted to high vowels. While there are clear phonetic differences between low vowels and high vowels, the learning outcomes are not based on phonetic differences, because of the manipulation of the acoustic cues that neutralised the intrinsic length advantage of low vowels and increased the perceptibility of high vowels.

These results strengthen the argument for a theory of phonology that is neither completely driven by phonetic changes over time, as held by Evolutionary Phonology (Blevins 2004), nor one that is based purely on phonetics (Ohala 2005). Evolutionary Phonology, for example, states that 'recurrent synchronic sound patterns have their origins in recurrent phonetically motivated sound change' (Blevins 2004: 8). Under Evolutionary Phonology thinking, the preference for stress to occur on low vowels over high vowels results from two sources: the phonetic features of

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

low vowels and learned phonological representations based on transmitted speech. The pertinent phonetic feature of low vowels is duration. Specifically, low vowels have greater duration than high vowels and have more acoustic energy, which fact then increases perception. Further, Evolutionary Phonology holds that sound change is passed on historically, from one generation to the next, and 'since every individual will have slightly different early childhood experiences, every individual will, by definition form a grammar based on distinct sets of surface forms' (Blevins 2004: 8). In discussing child language acquisition, Blevins indicates that 'innate phonological constructs' could include 'the prosodic units which function as the domains for stress and intonation contours' (Blevins 2004: 231). Although Blevins is not explicit, it is likely that she is referring to an innate knowledge of the syllable as a locus for stress. So to acquire the knowledge that stress is attracted to low vowels an individual would have to combine the phonetic knowledge of the relationship between duration and perceptibility *and* would have to have heard examples of stress preferring low vowels. Thus Evolutionary Phonology predicts that without previous experience in a phonological pattern there is no principled reason to choose that which is 'natural' over that which is 'unnatural'.[9]

If acquisition is based on what is perceived, then learners should acquire what they hear; there should be no difference between two patterns that are formally the same. Neither English nor French speakers have experience with stress being assigned on the basis of vowel height. Further, given the equalised duration of low and high vowels in the natural and unnatural experiments, there should be no significant difference in learning between the two systems. Therefore, the results reported here of an advantage for the low-vowel language cannot be due to any simple auditory prominence for those vowels. Although it might be argued that English speakers have some weak association between vowel height and stress, the same cannot be said for native French speakers. In addition to having fixed, predictable stress where vowel length or height is inconsequential, French speakers exhibit symptoms of being insensitive to stress in general. Without some pre-existing distinguishing mechanism for what is natural in language, there was no reason for them to learn one language better than the other, which they did. French and English speakers' preference for the natural language strongly suggests knowledge of phonological universals of the sonority hierarchy.

The advantage in learning the natural version demonstrates, I suggest, the *addition* of the operation of an innate language-learning mechanism to a general cognitive one. The exact form of the general cognitive mechanism is unclear and much research needs to be done to see

[9] Evolutionary Phonology makes no distinction between 'natural' and 'unnatural' in the sense that those terms have been used here. Both natural and unnatural sound patterns are viewed as products of historical sound changes (Blevins 2004: 81).

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

how cognition in non-linguistic areas intersects with language acquisition. Much of the work done so far in this area involves statistical learning. Research has shown that infants are able to use statistical regularities to segment a speech stream and discriminate the phonotactics of the ambient language (Saffran *et al.* 1996, Maye *et al.* 2002, Chambers *et al.* 2003). Adults also can detect and use statistical regularities in language to facilitate learning in an artificial language-learning situation (Mintz 2002). However, statistical learning would not have favoured one language over another in the current experiment, because the training and novel words were matched for consonant sequences and vowel quality across both languages. Thus statistical frequency of segments and transitional probabilities of syllables were the same for both languages.

Although statistical learning is not a good candidate for the cognitive mechanism at work in this experiment, there is evidence for other types of pattern detection. The psychology literature has much to say about pattern recognition in diverse areas, such as face detection (see Maurer 1985 for a summary of the literature review in this area) and number detection (e.g. Buckley & Gilman 1974, Wolters *et al.* 1987). Thus some cognitive mechanism could keep general track of how often stress occurs on words by syllable position, whether initial, medial or final. At the same time, or perhaps sequentially, a language-specific mechanism would have access to those patterns and, further, would be able to match these detected patterns to already existing, built-in biases towards preferred features of language.

Interaction of general cognitive mechanisms and an innate language-specific mechanism could take the following form: on being exposed to the linguistic input, the learner extracts one or more patterns from the input, in this case, recognising that stress occurs most often on the first syllable in the word and that it also occurs on some particular set of vowels. For the natural language group, their general pattern detection system enables them to recognise the pattern of stress occurring most often on the first syllable and on [ɑ] and [æ]. The particular phonemes mean nothing to the general cognitive system, but the language-specific mechanism, operating on the incoming linguistic data, recognises that there is something special about [ɑ] and [æ]: that those particular phonemes are low vowels, and thus good sites for stress attraction. In this way, the language-specific mechanism comes to the aid of the natural language group to identify stress attractors in the input, thus enabling them to be more successful in learning the pattern of stress occurring on low vowels. For the unnatural language group, recognising that the pattern of stress occurs most often on initial syllables as well as on [i] and [u] could similarly be explained through the use of general cognitive mechanisms. However, the general cognitive mechanism attributes no special feature to the phonemes [i] and [u], and any attempt to evaluate this pattern linguistically through use of the language-specific mechanism would fail, since this is not an attested linguistic pattern. The weight of learning the stress rule then falls solely

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

on general cognition, without the 'leg up' of the language-specific mechanism.

French learners of the natural rule provide some evidence of these hypothetical scenarios. However, even stronger support is provided by English learners of the natural language. French learners, with scores of 60%, 55% and 61%, were more accurate in correctly choosing words with initial and final stress over medial stress, but not significantly so. The score for final stress could be a reflection of their native stress pattern, in that they are more attuned to final stress than to stress on non-final syllables. In spite of this native tendency, however, they were able to move away somewhat from the stress pattern of their native language, as seen by their score of 60% on initial stress words. This shift towards choosing initial stress could also be argued as being a result of the general cognitive mechanism that allowed them to keep statistical track of the frequency of words with initial stress, along with some effect of the language-specific mechanism. This effect is suggested by the flat French scores of 55%, 55% and 55% on initial, medial and final stressed syllables of the unnatural language.

The results of the scores by syllable position among English learners more strongly support the learning scenarios described above. The natural language group's scores on initial, medial and final stress position was 70%, 73% and 69%. This means that when presented with a correct and an incorrect choice pair, they picked the correct choice most often, regardless of the stress position. The natural group thus demonstrated that they were close to fully understanding both the default and the main stress rules. The unnatural language group, with scores of 68%, 57% and 56% on initial, medial and final stress, was significantly more inconsistent. This inconsistency suggests that they were able to use their general cognitive mechanism to figure out that stress occurred most often on the initial syllable (the default rule) but, with the naturalness bias of the language-specific mechanism going against the stress pattern they heard, it was more difficult for them to figure out the main stress rule, as it applied to medial- and final-syllable stress.

These results add to the growing body of experimental evidence in support of a bias towards naturalness (e.g. Wilson 2003, Pater & Tessier 2005, Moreton 2008). Learners' ability to do better on the natural language also suggests a psychological reality for linguistic constraints based on the Stressed Peak hierarchy, which details a preference for higher sonority, stressed-syllable peaks.

The difference in learning between the natural and unnatural versions demonstrates the difference between a general mechanism aided by an innate, language-specific mechanism, on the one hand, and a more general cognitive mechanism that is not specific to language, on the other. The experimental methodology captures the slight advantage provided to learners by the language-specific mechanism providing the learner with a linguistic 'leg up' into the language, so to speak.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

6.1 Theoretical implications and future direction

There are a number of positions on the role of an innate language-specific mechanism and language acquisition. If, as we claim, naturalness biases exist in a language-specific mechanism in order to aid learners in language acquisition, then where does general cognition end and language bias begin? One extreme view is that language is wholly a product of domain-general mechanisms, with no language-specific mechanism whatsoever (e.g. Christiansen & Chater 2008). Another extreme view is that Universal Grammar is solely responsible for all things linguistic with language operating as a separate organism or instinct within the brain (e.g. Chomsky 1965, Pinker 1994). More intermediate views present a case for both kinds of mechanisms (e.g. Zuraw 2007, Barrett *et al.* 2008, Moreton 2008). There seems no doubt that language learners are able to use non-language-specific mechanisms in the language acquisition process. This has been shown, for example, in the areas of statistical learning (Saffran *et al.* 1996, Maye *et al.* 2002) and over-regularisation of inconsistent linguistic input by both children and adults (Hudson Kam & Newport 2005, 2009). I have argued for an intermediate position, wherein there is an interaction of a general and a language-specific cognitive mechanism. In unnatural language learning, general cognitive mechanisms can perform functions such as keeping track of an operation on a particular syllable by position. The experimental method of pitting against each other two formally similar rules with only a difference in naturalness should be a useful tool to aid in better understanding our overall cognitive mechanisms, both general and language-specific.

One such study is underway at present, using the generalisation that stress is attracted to heavy syllables over light syllables, all things being equal. Another method of probing the difference between general cognition and a language-specific one would be to test children on the same question. Previous research has shown that adults bring certain cognitive advantages to second language learning, while children have an innate language-learning capability that usually outperforms adults (Olson & Samuels 1973, Snow & Hoefnagel-Höhle 1982, Hansen 1995). Testing children to see what results they produce on the same experiment could provide insight into the balance between general and language-specific cognition.

Studies of this type, which use natural/unnatural dichotomies, provide a rich source of future research to tease apart explanations of language-specific phenomena from domain-general ones. More research into other natural *vs.* unnatural systems needs to be done to add weight to this claim, but, at a minimum, these results support the idea that adult learners have some ability to access universal linguistic principles to aid them in acquiring a natural phonological pattern.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Appendix A

1 Training/familiarisation words

	natural language	*unnatural language*	*consonant sequence*
Training words 1	'badaka	'buduku	bdk
	'gikusi	'gækasæ	gks
	zu'pæta	za'pitu	zpt
	tigu'sa	tæga'su	tgs
Training words 2	'zakipizu	'zukæpæza	zkpz
	'kæsuzaga	'kisazugu	kszg
	tu'gabupa	ta'gubapu	tgbp
	ziki'pasæ	zækæ'pusi	zkps
	digitu'za	dægæta'zu	dgtz
Training words 3	'pætibu	'pitæba	ptb
	'tægika	'tigæki	tgk
	ki'bapi	kæ'bupæ	kbp
	budu'gæ	bada'gi	bdg
Training words 4	'datibisæ	'dutæbæsi	dtbs
	'bakabæzu	'bukubiza	bkbz
	si'tækaza	sæ'tikuzu	stkz
	kubus'sæka	kaba'siku	kbsk
	situgu'zæ	sætaga'zi	stgz
Training words 5	'zægasi	'zigusæ	zgs
	'pæsupa	'pisapu	psp
	ti'tazu	tæ'tuza	ttz
	gizu'da	gæza'du	gzd
Training words 6	'pibizudu	'pæbæzada	pbzd
	'gæsatiki	'gisutækæ	gstk
	pu'tækigu	pa'tikæga	ptkg
	dibi'bazu	dæbæ'buza	dbbz
	bididu'pa	bædæda'pu	bddp

2 Novel words: set 1

natural language	*unnatural language*	*consonant sequence*	*natural language*	*unnatural language*	*consonant sequence*
'dugizi	'dagæzæ	dgz	'kizisubi	'kæzæsabæ	kzsb
pi'bæza	pæ'bizu	pbz	'gæzaku	'gizuka	gzk
sudu'sæpa	sada'sipu	sdsp	puzu'sa	paza'su	pzs
'tadagæza	'tidugizu	tdgz	'bækagæ	'bikugi	bkg
pi'gætu	pæ'gita	pgt	ki'kazæ	kæ'kuzi	kkz
'kazætæsi	'kuzitisæ	kzts	pitu'za	pæta'zu	ptz
buguzi'gæ	bagazæ'gi	bgzg	kuputi'da	kapatæ'du	kptd
'kæbigu	'kibæga	kbg	gu'sætæki	ga'sitikæ	gstk
guti'basu	gatæ'busa	gtbs	'pazukætu	'puzakita	pzkt

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

3 Novel words: set 2

natural language	*unnatural language*	*consonant sequence*	*natural language*	*unnatural language*	*consonant sequence*
'bizidu	'bæzæda	bzd	'kætabugæ	'kitubagi	ktbg
'gasatæ	'gusuti	gst	'sædipaku	'sidæpuka	sidpk
'sædaka	'siduku	sdk	'sagikæta	'sugækitu	sgkt
'sudisi	'sadæsæ	sds	'sazupida	'suzapædu	szpd
'dapiku	'dupæka	dpk	'tæbukuga	'tibakagu	tbkg
'kækibi	'kikæbæ	kkb	'zæbadusæ	'zibudasi	zbds
'ditizi	'dætæzæ	dtz	'puditiki	'padætækæ	pdtk
'sæpatu	'siputæ	spt	'gætubæbæ	'gitabibi	gtbb
'sapæbu	'supiba	spb	'zupitisi	'zapætæsæ	zpts
'dædugu	'didaga	ddg	'tapagiki	'tupugækæ	tpgk
'pagida	'pugædu	pgd	'pæsadazæ	'pisuduzi	psdz
'gætusæ	'gitasi	fts	'kapatapa	'kuputupu	kptp
bu'zata	bæ'zutu	bzt	zi'batikæ	zæ'butæki	zbtk
ti'gæti	tæ'gitæ	tgt	du'sæbutu	da'sibata	dsbt
di'paga	dæ'pugu	dpg	pi'pæbatu	pæ'pibuta	ppbt
bi'gadu	bæ'guda	bgd	su'pæsaba	sa'pisubu	spsb
su'tasi	sa'tusæ	sts	piku'bætu	pæka'bita	pkbt
zi'kæda	zæ'kidu	zkd	supu'dæka	sapa'diku	spdk
kuki'pæ	kakæ'pi	kkp	dizi'bæpa	dæzæ'bipu	dzbp
zupi'zæ	zapæ'zi	zpz	bipu'bapi	bæpa'bupa	bpbp
tugu'za	taga'zu	tgz	dudizu'pa	dadæza'pu	ddzp
pigi'sa	pægæ'su	pgs	sigibi'tæ	sægæbæ'ti	sgbt
tuki'ka	takæ'ku	tkk	gutuki'da	gatakæ'du	gtkd
duzu'dæ	daza'di	dzd	zibugu'ka	zæbaga'ku	zbgk

Appendix B: Participant instructions

1 *Written instructions given at the beginning of the experiment*

You are going to hear words in a language that you've never heard before. Then you'll be tested to see how well you can identify words that belong to that language. You will hear words with stress on different syllables. Stress can be described as the part of the word that sounds more prominent than another part. In English we can usually say exactly which syllable in a word should be more prominent than the others. For example, we say *banána*, not *bananá*. We can usually even decide which syllable should be more prominent in nonsense words, such as *poedéktal* and not *póedektal*. In this new language you will learn to do the same thing. After hearing several repetitions of words in the language, you will hear pairs of words that differ in their stress pattern. You will be asked to choose which word sounds like it belongs to the language you are learning.

First I'll give you a practice so you can get familiar with the process.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

2 *After the practice* (*the AXB pre-training*)

If you feel comfortable with how the experiment works, we can now begin. First, just listen to the words of the new language. Then you will be asked to choose which words belong. Press the spacebar to advance within the experiment.

3 *Experiment overview* (*presented on the computer monitor*)

Each section of this experiment has a learning portion and a testing portion. Each word is accompanied by a picture of an object. In the learning sections listen to the words carefully and try to learn them. In the testing portion you will be given two words that are the same except that the stress is on a different syllable. You are to choose which word is the one you learned by pressing the RED key (1) or the BLUE key (2).

First you will listen to a group of 4 different words repeated 4 times, and then you'll be tested on those 4. Next you will learn 5 new words, also repeated 4 times, and be tested on those 5. Then you'll be tested on the first 9 words.

The process will be repeated with 9 new words, where you'll learn and be tested on 4, then learn and tested and 5, and then tested on those 9. Following that section you will be tested on the 18 words learned so far plus some new words you didn't have in any of the previous sections.

After learning 9 more new words (a total of 27 words) and being tested on all 27 words, you will have a final test with novel words that you have not previously heard.

Press the spacebar to begin.

4 *After the end of each training section the instructions for the testing of those words were as follows:*

You have just finished the first (second, third, etc.) training section. You'll now be tested on the words you just heard.

You'll hear groups of two words. Both words are the same, but with different stress. Pick the word that sounds like one you heard before. After you choose you will hear the correct answer.

Press the RED key (1) if the 1st word sounds right, or press the BLUE key (2) if the 2nd word sounds right.

Press the spacebar to begin.

5 *At the start of the final test with novel words*

You have reached the final test. You'll hear 48 new words that have the same stress pattern as the ones you have learned.

Once again, you'll be given pairs of words that are the same except that stress is on a different syllable. Choose the word that sounds as if it belongs to the language.

Press the RED key (1) if the 1st word sounds right, or press the BLUE key (2) if the 2nd word sounds correct.

Press the spacebar to continue.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

REFERENCES

Abramson, Marianne & Stephen D. Goldinger (1997). What the reader's eye tells the mind's ear: silent reading activates inner speech. *Perception and Psychophysics* **59**. 1059–1068.

Aleman, André & Mascha van 't Wout (2004). Subvocalization in auditory-verbal imagery: just a form of motor imagery? *Cognitive Processing* **5**. 228–231.

Archibald, John (1995). The acquisition of stress. In John Archibald (ed.) *Phonological acquisition and phonological theory*. Hillsdale, NJ: Erlbaum. 81–109.

Armstrong, Susan (1999). *Stress and weight in Quebec French*. MA thesis, University of Calgary.

Baddeley, Alan, Marge Eldridge & Vivien Lewis (1981). The role of subvocalisation in reading. *Quarterly Journal of Experimental Psychology* **33A**. 439–454.

Barrett, H. Clark, Willem E. Frankenhuis & Andreas Wilke (2008). Adaptation to moving targets: culture/gene coevolution, not either/or. *Behavioral and Brain Sciences* **31**. 511–512.

Blevins, Juliette (2004). *Evolutionary Phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.

Boersma, Paul (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* **21**. 43–58.

Boersma, Paul & David Weenink (2006). *Praat: doing phonetics by computer* (version 4.5.02). http://www.praat.org/.

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, P. B. & C. B. Gilman (1974). Comparison of digits and dot patterns. *Journal of Experimental Psychology: General* **103**. 1131–1136.

Burzio, Luigi (1994). *Principles of English stress*. Cambridge: Cambridge University Press.

Cardona, George (1965). *A Gujarati reference grammar*. Philadelphia: University of Pennsylvania Press.

Chambers, Kyle E., Kristine H. Onishi & Cynthia Fisher (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition* **87**. B69–B77.

Charette, Monik (1991). *Conditions on phonological government*. Cambridge: Cambridge University Press.

Chomsky, Noam (1965). *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

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

Christiansen, Morten H. & Nick Chater (2008). Language as shaped by the brain. *Behavioral and Brain Sciences* **31**. 489–509.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In John Kingston & Mary E. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press. 283–333.

Cygankin, D. V. & C. Z. Davaev (1975). *Ocherk sravnitel'oi grammatiki mordovskix* (*mokshanskogo i erz'anskogo*) *literaturnyx iazykov*. Saransk: Mordovskii Gosudarstvennyi Universitet Imeni N. P. Ogareva.

Dechaine, Rose-Marie (1990). Stress and weight gain. Ms, University of Massachusetts, Amherst.

de Lacy, Paul (2002). *The formal expression of markedness*. PhD dissertation, University of Massachusetts, Amherst.

de Lacy, Paul (2004). Markedness conflation in Optimality Theory. *Phonology* **21**. 145–199.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Delattre, Pierre (1939). Accent de mot et accent de groupe. *The French Review* **13**. 141–146.
Dell, François (1980). *Generative phonology and French phonology*. Cambridge: Cambridge University Press.
Dell, Gary S., Kristopher D. Reed, David R. Adams & Antje S. Meyer (2000). Speech errors, phonotactic constraints, and implicit learning: a study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **26**. 1355–1367.
Dupoux, Emmanuel, Christophe Pallier, Núria Sebastián & Jacques Mehler (1997). A destressing 'deafness' in French? *Journal of Memory and Language* **36**. 406–421.
Esper, Erwin Allen (1925). *A technique for the experimental investigation of associative interference in Artificial Language Material*. Philadelphia: Linguistic Society of America.
Fant, Gunnar, Anita Kruckenberg & Lennart Nord (1991). Durational correlates of stress in Swedish, French and English. *JPh* **19**. 351–365.
Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *JASA* **27**. 765–768.
Gordon, Matthew (2002). A phonetically driven account of syllable weight. *Lg* **78**. 51–80.
Halle, Morris & Jean-Roger Vergnaud (1987). Stress and the cycle. *LI* **18**. 45–84.
Hammond, Michael (1999). *The phonology of English: a prosodic optimality-theoretic approach*. Oxford: Oxford University Press.
Hansen, Doris (1995). A study of the effect of the acculturation model on second language acquisition. In Fred R. Eckman, Diane Highland, Peter W. Lee, Jean Milcham & Rita Rutkowski Weber (eds.) *Second language acquisition theory and pedagogy*. Mahwah, NJ: Erlbaum. 305–316.
Hayes, Bruce (1982). Extrametricality and English stress. *LI* **13**. 227–276.
Hayes, Bruce (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.
Hayes-Harb, Rachel, Janet Nicol & Jason Barker (2010). Learning the phonological forms of new words: effects of orthographic and auditory input. *Language and Speech* **53**. 367–381.
Healy, Alice F. & Andrea G. Levitt (1980). Accessibility of the voicing distinction for learning phonological rules. *Memory and Cognition* **8**. 107–114.
Helimski, Eugene (1998). Nganasan. In Daniel Abondolo (ed.) *The Uralic languages*. London & New York: Routledge. 480–515.
Hubbard, Timothy L. (2010). Auditory imagery: empirical findings. *Psychological Bulletin* **136**. 302–329.
Hudson Kam, Carla L. & Elissa L. Newport (2005). Regularizing unpredictable variation: the roles of adult and child learners in language formation and change. *Language Learning and Development* **1**. 151–195.
Hudson Kam, Carla L. & Elissa L. Newport (2009). Getting it right by getting it wrong: when learners change languages. *Cognitive Psychology* **59**. 30–66.
Itkonen, E. (1955). Über die Betonungsverhältnisse in den finnisch-ugrischen Sprachen. *Acta Linguistica Academiae Scientiarum Hungaricae* **5**. 21–34.
Jassem, Wiktor, John Morton & Maria Steffen-Batog (1968). The perception of stress in synthetic speech-like stimuli by Polish listeners. In Wiktor Jassem (ed.) *Speech analysis and synthesis*. Vol. 1. Warsaw: Państwowe Wydawnictwo Naukowe. 289–308.
Jespersen, Otto (1904). *Lehrbuch der Phonetik*. Leipzig & Berlin: Teubner.
Kager, René (1989). *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.
Kenstowicz, Michael (1994). Sonority-driven stress. Ms, MIT. Available as ROA-33 from the Rutgers Optimality Archive.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Kiparsky, Paul (1979). Metrical structure assignment is cyclic. *LI* **10**. 421–441.
Krause, Scott (1980). *Topics in Chukchee phonology and morphology*. PhD dissertation, University of Illinois, Urbana-Champaign.
Ladefoged, Peter (2003). *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Malden, Mass. & Oxford: Blackwell.
McCarthy, John J. (2002). *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.
McCarthy, John J. (2003). OT constraints are categorical. *Phonology* **20**. 75–138.
McCarthy, John J. & Alan Prince (1993). Generalized alignment. *Yearbook of Morphology 1993*. 79–153.
McCarthy, John J. & Alan Prince (1995). Faithfulness and reduplicative identity. In Jill N. Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*. Amherst: GLSA. 249–384.
Mairs, Jane Lowenstein (1989). Stress assignment in interlanguage phonology. In Susan M. Gass & Jacquelyn Schachter (eds.) *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press. 260–283.
Martin, Pierre (2002). Le système vocalique du français du Québec: de l'acoustique à la phonologie. *La linguistique* **38:2**. 71–88.
Maurer, Daphne (1985). Infants' perception of facedness. In Tiffany M. Field & Y. Nathan Fox (eds.) *Social perception in infants*. Norwood, NJ: Ablex. 73–100.
Maye, Jessica, Janet F. Werker & LouAnn Gerken (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82**. B101–B111.
Mintz, Toben H. (2002). Category induction from distributional cues in an artificial language. *Memory and Cognition* **30**. 678–686.
Moreton, Elliott (2008). Analytic bias and phonological typology. *Phonology* **25**. 83–127.
Ohala, John J. (2005). Phonetic explanations for sound patterns: implications for grammars of competence. In William J. Hardcastle & Janet Mackenzie Beck (eds.) *A figure of speech: a Festschrift for John Laver*. Mahwah, NJ: Erlbaum. 23–38.
Olson, Linda L. & S. Jay Samuels (1973). The relationship between age and accuracy of foreign language pronunciation. *Journal of Educational Research* **66**. 263–268.
Onishi, Kristine H., Kyle E. Chambers & Cynthia Fisher (2002). Learning phonotactic constraints from brief auditory experience. *Cognition* **83**. B13–B23.
Paradis, Claude & Denise Deshaies (1990). Rules of stress assignment in Québec French. *Language Variation and Change* **2**. 135–154.
Pater, Joe (1992). The acquisition of parameters for word stress by French learners of English. Ms, Concordia University.
Pater, Joe (1997). Metrical parameter missetting in second language acquisition. In S. J. Hannahs & Martha Young-Scholten (eds.) *Focus on phonological acquisition*. Amsterdam & Philadelphia: Benjamins. 235–261.
Pater, Joe (2000). Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints. *Phonology* **17**. 237–274.
Pater, Joe & Anne-Michelle Tessier (2005). Phonotactics and alternations: testing the connection with artificial language learning. In Katherine Flack & Shigeto Kawahara (eds.) *UMOP 31: Papers in experimental phonetics and phonology*. Amherst: GLSA. 1–16.
Payne, Judith (1990). Asheninca stress patterns. In Doris L. Payne (ed.) *Amazonian linguistics: studies in lowland South American languages*. Austin: University of Texas Press. 185–209.
Peperkamp, Sharon & Emmanuel Dupoux (2002). A typological study of stress 'deafness'. In Carlos Gussenhoven & Natasha Warner (eds.) *Laboratory Phonology 7*. Berlin & New York: Mouton de Gruyter. 203–240.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Peperkamp, Sharon & Emmanuel Dupoux (2007). Learning the mapping from surface to underlying representations in an artificial language. In Jennifer Cole & Jose Ignacio Hualde (eds.) *Laboratory Phonology 9*. Berlin & New York: Mouton de Gruyter. 315–338.

Peperkamp, Sharon, Emmanuel Dupoux & Núria Sebastián-Gallés (1999). Perception of stress by French, Spanish and bilingual subjects. Paper presented at 7th European Conference on Speech Communication and Technology (Eurospeech), Aalborg.

Peperkamp, Sharon, Katrin Skoruppa & Emmanuel Dupoux (2006). The role of phonetic naturalness in phonological rule acquisition. In David Bamman, Tatiana Magnitskaia & Colleen Zaller (eds.) *Proceedings of the 30th Annual Boston University Conference on Language Development*. Somerville: Cascadilla. 464–475.

Pinker, Steven (1994). *The language instinct*. New York: William Morrow.

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.

Pycha, Anne, Pawel Nowak, Eurie Shin & Ryan Shosted (2003). Phonological rule-learning and its implications for a theory of vowel harmony. *WCCFL* **22**. 423–435.

Rice, Curt (1996). Apparent exceptional penultimate stress in English. *Nordlyd* **24**. 157–167.

Saffran, Jenny R., Richard N. Aslin & Elissa L. Newport (1996). Statistical learning by 8-month-old infants. *Science* **274**. 1926–1928.

Saffran, Jenny R. & Erik D. Thiessen (2003). Pattern induction by infant language learners. *Developmental Psychology* **39**. 484–494.

Schane, Sanford A. (1968). *French phonology and morphology*. Cambridge, Mass.: MIT Press.

Schane, Sanford A., Bernard Tranel & Harlan Lane (1974). On the psychological reality of a natural rule of syllable structure. *Cognition* **3**. 351–358.

Sievers, Eduard (1881). *Grundzüge der Phonetik*. Leipzig: Breitkopf & Härtel.

Smith, Jennifer L. (2002). *Phonological augmentation in prominent positions*. PhD dissertation, University of Massachusetts, Amherst.

Snow, Catherine E. & Marian Hoefnagel-Höhle (1982). The Critical Period for language acquisition: evidence from second language learning. In Stephen Krashen, Robin Scarcella & Michael Long (eds.) *Child–adult differences in second language acquisition*. Rowley, MA: Newbury House. 93–111.

Stampe, David (1979). *A dissertation on Natural Phonology*. New York: Garland.

Tesar, Bruce & Paul Smolensky (1998). Learnability in Optimality Theory. *LI* **29**. 229–268.

Tesar, Bruce & Paul Smolensky (2000). *Learnability in Optimality Theory*. Cambridge, Mass.: MIT Press.

Thorn, Annabel S. C. & Susan E. Gathercole (2001). Language differences in verbal short-term memory do not exclusively originate in the process of subvocal rehearsal. *Psychonomic Bulletin and Review* **8**. 357–364.

Tranel, Bernard (1987). *The sounds of French : an introduction*. Cambridge: Cambridge University Press.

Vaux, Bert (1998). *The phonology of Armenian*. Oxford: Clarendon Press.

Walker, Douglas (2001). *French sound structure*. Calgary: University of Calgary Press.

Watson, C. S. (1973). Psychophysics. In Benjamin B. Wolman (ed.) *Handbook of general psychology*. Englewood Cliffs: Prentice Hall. 275–306.

Wilson, Colin (2003). Experimental investigation of phonological naturalness. *WCCFL* **22**. 533–546.

Wilson, Colin (2006). Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* **30**. 945–982.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press

Wolters, Gezinus, Hanneke van Kempen & Gert-Jan Wijlhuizen (1987). Quantification of small numbers of dots: subitizing or pattern recognition? *The American Journal of Psychology* **100**. 225–237.
Yallop, Colin (1977). *Alyawarra: an Aboriginal language of Central Australia*. Canberra: Australian Institute of Aboriginal Studies.
Zhang, Jie & Yuwen Lai (2010). Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology* **27**. 153–201.
Zuraw, Kie (2007). The role of phonetic knowledge in phonological patterning: corpus and survey evidence from Tagalog infixation. *Lg* **83**. 277–316.

https://doi.org/10.1017/S0952675710000199 Published online by Cambridge University Press