It is easier to learn the meaning of forms with a canonical stress pattern

Robert Daland & Yun Jung Kim

UCLA Dep't of Linguistics
r.daland@gmail.com, yun.ucla@gmail.com

ABSTRACT

Generative linguistics is predicated on a *conceptual* distinction between the lexicon and the grammar. However in practice, lexical and grammatical acquisition are interdependent: phonotactics are projected in part from the lexicon [17], while the underlying form of a morpheme depends on the pattern of contrast and neutralization determined by the grammar [25]. This paper demonstrates a causative effect of phonotactic knowledge on word-learning.
Stress position is contrastive in English, but trochees are statistically predominant in the lexicon. This study asks whether it is easier to learn novel trochaic words than iambic ones. It was found that novel trochaic and iambic *forms* are remembered equally well, but the form-meaning association was better learned for trochees.

Keywords: word learning; phonotactics; stress

1. INTRODUCTION

Word-learning plays a central role in language acquisition. Aside from representing the entities and events that we talk about, words also provide an important seat for grammatical generalizations. For example, a consensus position is that at least some phonotactic knowledge is projected from the lexicon [2**,** 4**,** 17]. Put simply, word-learning shapes and constrains grammar learning.

Grammar learning also shapes word-learning. For example, French stress is assigned at or near the end of a phonological phrase, regardless of what or how many words the phrase contains. Therefore, it is widely agreed that stress information is simply not a part of French lexical representations, and therefore does not need to be learned during word learning [16]. Properties which nearly always signal a meaning difference, such as obstruent voicing in French, must be stored in lexical representations [25].

The position of English stress is lexically contrastive, e.g. *REcord* is an object containing some data about an event or person, while *reCORD* refers to the action of making such an object. However, English stress is unlike French obstruent voicing in that it is largely predictable on the basis of other information [10]. For example, [3] found that approximately 75% of English dictionary entries are stress-initial, and it is likely that the proportion is much higher in the disyllables learned during early adolescence [23]. English-learning infants have formed the generalization that words tend to be trochaic as early as 9 months of age, as evident from deploying this generalization for word segmentation [13]. Thus, English stress exhibits a statistical default, but a minority of items contrast with the default.

There is a long-running debate as to the content of lexical representations in such cases. For example, one standard position is that items with the default pattern do not contain any lexical specification of the relevant property, and it is derived by the grammar, while items with a non-default pattern must have it listed in the lexical representation [18]. We do not address the content of lexical representations directly here. Rather, we take up the more empirically accessible question of whether words with a default stress pattern are more easily *learned* than those with a non-default stress pattern.

2. PREVIOUS WORK

2.1. Learning the grammar of stress

Constraint-based phonology has ushered in a new era for phonological acquisition. Work by Tesar and coleagues has illustrated how many grammatical generalizations can be acquired from surface evidence alone [25]. In the domain of stress learning, constraints have been proposed to generate a broad subset of all attested stress patterns [7**,** 14]. Recent work has addressed the learning of stress patterns, a difficult problem owing to 'hidden structure' -- metrical units such as feet that are believed not to be overt in the signal [11**,** 12**,** 20]. It is not clear how relevant most of this work is for English, since it focuses only on languages with predictable stress. Fortunately, what matters here is the simple generalization that English prefers trochees over iambs, and that is easy to learn [23].

2.2. Word learning

Recent years have seen a surge of research on word-learning [25]. We have learned that sleep facilitates lexicalization [5], and that the nature of the exposure and subsequent usage strongly conditions properties of the lexical representation [15]. Of greater relevance for the present study, there have been a number of studies on how specific linguistic properties of a word's form affect the ease with which it is learned.

Perhaps the most well-known result is that learning a novel form-meaning association is difficult for young children if the form is very similar to a familiar word (e.g. *ball* blocks *gall*, [25]). Prior exposure to the form in a segmentation task supports word-learning [8], and Fennell has has argued convincingly for a resource limitations account (e.g. [6]). Thus, one might be tempted to hypothesize that early word-learners are at an overall disadvantage in learning words that are similar to familiar forms.

This hypothesis is not consistent with the results of a series of studies by Storkel and colleagues, who Invoking the conceptual distinction between *phonotactic typicality* and *nonword density*, [21**, 22**]. A nonword is phonotactically typical if it is composed of high-frequency subsequences, and it is in a high-density neighborhood if there are several existing words differing only in a single sound (typical-high: [pim], atypical-low: [waf], typical-low: [han], atypical-high [hif]). Aside from the expected effect of exposure frequency, [21] found a facilitatory effect of neighborhood density, but an inhibitory effect of phonotactic typicality in adult learners. An analogous dissociation is reported in [22] for early child learners.

In the case of English disyllables, trochaic stress is clearly more phonotactically typical. By analogy with [21], we might expect adult learners to be disadvantaged at learning form-meaning associations for the phonotactically typical pattern (trochaic stress) relative to the atypical (iambic stress) pattern. That is the prediction tested here.

3. METHODS

3.1. Participants

Native speakers of English (*n*=22) were recruited from the UCLA Linguistics/Psychology subject pool. Language background was assessed by a detailed survey; speakers were not excluded if they had additional languages, but only if they were not exposed to English from birth.

3.2. Stimuli

The nonwords consisted of 20 CVCVC forms that were phonotactically acceptable with both trochaic and iambic stress, e.g. /pɛʃɪf/, trochaic --> [ˈpɛʃəf], iambic --> [pəˈʃɪf]. To distinguish whether any word-learning effects derive specifically from stress, or from phonotactic probability more generally, a phonotactic *z*-score was calculated. The log probability of a pronounced form was calculated as the average log probability assigned by syllabic bigram model and the syllabic parser of [2], which were the two best models for predicting well-formedness judgments of licit items in a recent 'bake-off' [4]. The phonotactic *z*-score of a form was calculated from the mean and deviation of a larger set of 240 CVCV and CVCVC nonwords developed in a pilot experiment.

Table 1: Target nonwords, with phonotactic *z*-scores.

UR	trochee	*z*_plog	iamb	*z*_plog
/fomɝk/	[ˈfomɚk]	1.71	[fəˈmɝk]	0.21
/fumos/	[ˈfumos]	0.53	[fəˈmos]	1.36
/gɛmɪf/	[ˈgɛməf]	0.16	[gəˈmɪf]	-2.01
/giʃʌb/	[ˈgiʃəb]	-0.50	[gɪˈʃʌb]	-2.15
/jalam/	[ˈjaləm]	0.51	[jəˈlam]	0.17
/d͡ʒɪbok/	[ˈd͡ʒɪbok]	-1.40	[d͡ʒɪˈbok]	0.05
/kɛpʊk/	[ˈkɛpək]	0.21	[kəˈpʊk]	1.09
/lɛbɪθ/	[ˈlɛbəθ]	-0.46	[ləˈbɪθ]	0.28
/lot͡ʃɪb/	[ˈlot͡ʃəb]	-0.02	[ləˈt͡ʃɪb]	-1.51
/mafɪm/	[ˈmafəm]	0.50	[məˈfɪm]	-0.72
/net͡ʃiv/	[ˈnet͡ʃəv]	0.30	[nəˈt͡ʃiv]	0.38
/nɝmɪb/	[ˈnɝməb]	-0.75	[nɚˈmɪb]	-0.46
/pɛʃɪf/	[ˈpɛʃəf]	-0.73	[pəˈʃɪf]	-0.32
/pogaf/	[ˈpogəf]	0.56	[pəˈgaf]	-1.14
/posuk/	[ˈposək]	0.91	[pəˈsuk]	-1.01
/ræsɪm/	[ˈræsɪm]	2.52	[rəˈsɪm]	1.90
/ʃɛpof/	[ˈʃɛpəf]	-0.26	[ʃəˈpof]	0.09
/teðin/	[ˈteðɪn]	0.61	[təˈðin]	-1.17
/vaθig/	[ˈvaθɪg]	-1.84	[vəˈθig]	-1.59
/zɛdug/	[ˈzɛdəg]	-1.32	[zəˈdug]	-0.50

The auditory stimuli were produced by a phonetically trained female native speaker of American English (Midwestern dialect). She produced the tokens in Table 1 with each stress pattern, including unstressed vowel reduction.

The visual stimuli consisted of a set of 'alien pictures' designed for word-learning studies [9].

Figure 1: Example of an alien picture.

3.3. Procedure

After completing the language background survey and consenting, participants were seated at a testing station, and put on sound-cancelling headphones. They completed a multi-stage word-learning and -testing protocol.

Exposure I. Listeners were exposed to nonword-picture pairing. Trials began with a fixation cross (300 ms) followed by the alien picture, which stayed on screen for 3 s. The training token was embedded in a carrier phrase and played twice, once at the onset of the visual stimulus, and once halfway through the trial (e.g. *This is a* [ˈgɛməf].. *This is a* [ˈgɛməf]). There were 2 blocks, with 1 trial for each alien/name pair per block in random order.

Nonword recall. Listeners were tested on their ability to remember the forms of exposed alien names. Trials began with a fixation cross, followed by auditory presentation of the test token. Listeners were instructured to press YES (keypad 's') if they had heard the nonword during Exposure I, and NO (keypad 'k') otherwise. All 20 exposed alien names and 20 prosodically-matched foils were presented in random order. (The foils were drawn from the same pool of CVCVC forms as the alien names, and were recorded in the same recording session.)

Nonword repetition. Listeners repeated alien names and foils out loud after auditory presentation. This production data is not analyzed here.

Exposure II. The procedure was similar to Exposure I, except that the name was presented just once, in isolation, and then the listener was asked to repeat the alien name (e.g. [ˈgɛməf].. _____); and there was only one block. This production data is not analyzed here.

Picture identification. Listeners were presented auditorily with an alien name, and were asked to select the corresponding alien picture from among 4 options. Trials began with a fixation cross (300 ms), after which four alien pictures were displayed in the four quadrants of the screen. Listeners pressed corner keys on the keypad to indicate their selection. All 20 alien names were presented once each. Each corner was the correct answer an equal number of times, and the occurrence of each alien picture in each position was counterbalanced.

Picture naming with feedback. Listeners were presented with an alien picture, and were asked to say the alien name aloud. After 5 seconds, the correct response was presented auditorily. All 20 alien pictures were presented once each. This production data is not analyzed here.

Picture naming. The procedure was the same as in the preceding set, except the correct name was not supplied. This production data is not analyzed here.

4. RESULTS AND DISCUSSION

4.1. Results

Nonword recall results. The nonword recall responses were subjected to a logistic mixed-effects regression using stress pattern as the fixed effect, with random intercepts for listener and item. Stress pattern was not significant (β=0.32, p>0.272). The percent familiar responses are shown in Fig. 2, as a function of stress pattern and actual familiarity.

Picture identification results. The picture identification responses were subjected to a logistic mixed-effects regression using stress pattern as the fixed effect, with random intercepts for listener. Stress pattern was highly significant (β=1.24, p<.001), with a greater rate of correct responses in the trochaic condition. The percent correct responses are shown for both conditions in Fig. 3.

The same analyses were conducted with the phonotactic z scores (scalar) instead of stress (binary). These results are not reported, since the phonotactic coefficient was not significantly different than 0 in either case, and the resulting models had considerably lower (poorer) log-likelihoods and BICs.

In summary, the form-level property of stress did not affect performance on nonword recall (a form-level task). However, stress did affect performance on picture identification, an index of the association between form and meaning.

Figure 2: Proportion correct on targets vs. foils. (NB Correct response is 'familiar' for a target, but 'unfamiliar' a foil.)

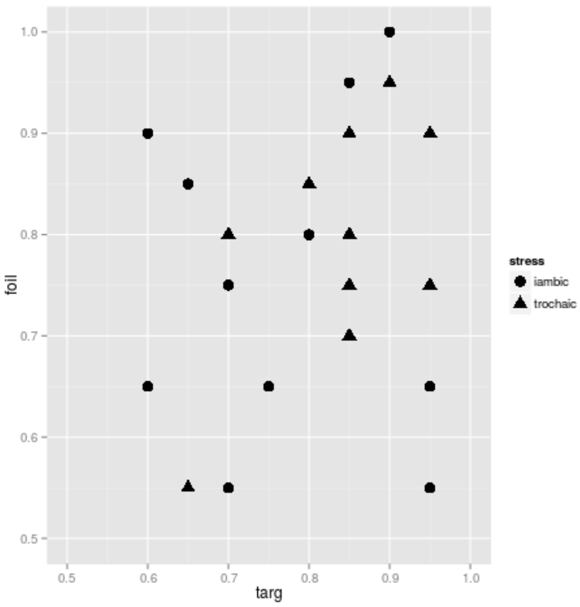
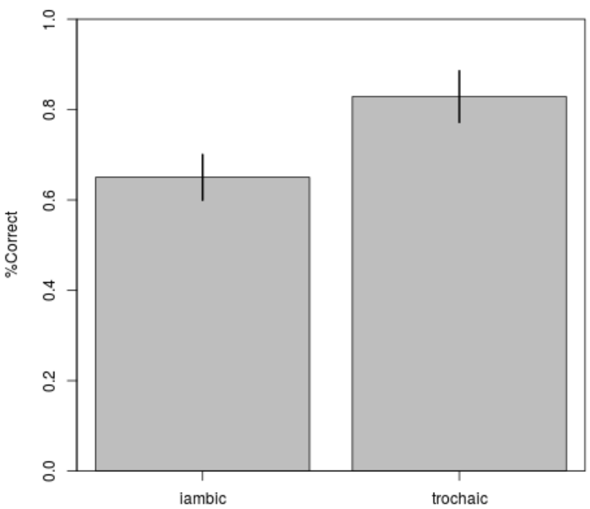


Figure 3: Percentage of correct picture identification in iambic vs. trochaic condition.



4.2. Discussion

These results complement and extend previous results on the interaction between word-learning and grammar learning throughout the lifespan. Previous work has shown that the complete absence of lexical stress results in an inability to remember lexical stress in online experiments [16]. Other work has considered how a learner might discover that a phonological property is lexically contrastive, although that has largely been at the level of segmental contrasts [25]. Still other work has considered the role of relative phonotactic typicality during word-learning, again mostly at the segmental level [21, 22]. However, comparatively little work has considered the acquisition of stress contrasts when there is a strong statistical tendency or grammatical default in the language. This is exactly the case with English, which strongly prefers trochaic stress in disyllables, but nevertheless allows iambic stress in a significant minority of forms.

The present study suggests that listeners are specifically advantaged at learning a novel form-meaning association when the form exhibits the grammatical default -- in this case, trochaic stress. No such advantage was seen in nonword recall.

This suggests that non-default phonological properties are not harder to encode *per se*. The 'phonological loop' that is thought to underlie the acquisition of novel wordforms is believed to rely on an auditory buffer with a memory on the order of a few seconds [1]. The phonological loop relies on phonological encoding of speech; subvocal 'rehearsal' of a novel form allows it to be retained as long as attention is not diverted to new stimuli. However, nonword recall cannot rely on short-term memory representations only, since the test occurs several minutes after the initial exposure, and after many intervening stimuli. It seems reasonable to suppose that nonword recall relies on some kind of intermediate-term memory of the phonologically encoded form [19].

The contrast between default and non-default items is only apparent in the picture identification task, which assesses the extent to which the form-meaning *association* has been learned. The most reasonable interpretation for this finding is that non-default phonological properties require additional cognitive resources to *bind* to a semantic representation, and that is why the binding process is evidently more successful for English trochaic disyllables than iambic ones. Though highly speculative, this interpretation is at least consistent with so-called 'underspecification' theories, which posit that only non-default phonological information is stored in lexical representations [16, 18]. An additional point in favor of this interpretation is the fact that the binary variable of stress predicted picture identification accuracy better than the gradient measure of phonotactic typicality. Representations are expected to contain binary features, whereas gradient effects are expected to fall out from processing.

7. REFERENCES

[1] Baddeley, A., Gathercole, S., Papagno, C. 1998. The phonological loop as a language learning device. *Psychological Review* 105(1), 158-73. doi:10.1037//0033-295X.105.1.158

[2] Coleman, J., Pierrehumbert, J. B. 1997. Stochastic phonological grammars and acceptability. *3rd Meeting of the ACL Special Interest Group in Computational Phonology: Proceedings of the Workshop, 12 July 1997*. Association for Computational Linguistics, Somerset NJ. 49-56.

[3] Cutler, A., Carter, D. M. 1987. The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language* 2, 133-142. DOI: 10.1016/0885-2308(87)90004-0

[4] Daland, R., Hayes, B., White, J., Garellek, M., Davis, A. Norrmann, I. 2011. Explaining sonority projection effects. *Phonology* 28(2), 197-234.

[5] Dumay, N., Gaskell, M. G. 2007. Sleep-associated changes in the mental representation of spoken words. *Psychological Science* 18, 35-39.

[6] Fennell, C.T. 2011. Object familiarity enhances infants' use of phonetic detail in novel words. *Infancy* 16(3), 1-15.

[7] Gordon, M. 2002. A factorial typology of quantity-insensitive stress. *Natural Language and Linguistic Theory* 20, 491-552.

[8] Graf Estes, K., Evans J. L., Alibali, M. W., Saffran J. R. 2007. Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science* 18(3), 254-60.

[9] Gupta, P., Lipinksi, J., Abbs, B., Lin, P. H., Aktunc E., Ludden, D., Martin, N., Newman, R. 2004. Space aliens and nonwords: stimuli for investigating the learning of novel word-meaning pairs. *Behavioral Research Methods, Instruments, and Computers* 36(4), 599-603.

[10] Hayes, B. 1982. Extrametricality and English stress. *Linguistic Inquiry* 13, 227-276.

[11] Heinz, J. 2009. On the role of locality in learning stress patterns. *Phonology* 26, 303-351. doi:10.1017/S0952675709990145

[12] Jarosz, G. 2013. Learning with hidden structure in Optimality Theory and Harmonic Grammar: Beyond Robust Interpretive Parsing. *Phonology* 30(1), 27-71.

[13] Jusczyk, P. W., Houston, D. M., Newsome, M. 1997. The beginnings of word segmentation in English-learning infants. *Cognitive Psychology* 39(3-4), 159-207.

[14] Kager, R. 2012. Stress in windows: Language typology and factorial typology. *Lingua* 122, 1454-1493.

[15] Leach, L., Samuel, A. G. 2007. Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology* 55(4), 306-353. doi: 10.1016/j.cogpsych.2007.01.001

[16] Peperkamp, S., Vendelin, I., Dupoux, E. 2010. Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics* 38, 422-430.

[17] Pierrehumbert, J. B. 2003 Phonetic diversity, statistical learning, and acquisition of phonology, *Language and Speech* 46(2-3), 115-154.

[18] Pinker, S., Ullman, M. T. 2002. The past-tense debate. *TRENDS in Cognitive Science* 6(11), 456-463.

[19] Rosenzweig, M. R., Bennett, E. L., Colombo, P. J., Lee, D. W., Serrano, P. A. 1993. Short-term, intermediate-term, and long-term memories. *Behavioural Brain Research* 57(2), 193-198. doi:10.1016/0166-4328(93)90135-D

[20] Staubs, R. 2014. *Computational modeling of learning biases in stress typology*. Doctoral dissertation, University of Massachusetts Amherst.

[21] Storkel, H. L., Armbruster, J., Hogan, T. P. 2006. Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research* 49, 1175-1192. doi: 10.1044/1092-4388(2006/085)

[22] Storkel, H. L. & Lee, S. Y. 2011. The independenteffects of phonotactic probability and neighborhood density on lexical acquisition by preschool children. *Language and Cognitive Processes* 26(2), 191-211.

[23] Swingley, D. 2005. Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology* 50, 86-132.

[24] Swingley, D., Aslin, R.N. 2007. Lexical competition in young children's word learning. Cognitive Psychology 54, 99-132

[25] Tesar, B. 2006. Faithful contrastive features in learning. *Cognitive Science* 30(5), 863-903.

[26] Hall, D.G., Waxman, S.R. (Eds.). 2004. *From many strands: Weaving a lexicon*. Cambridge: MIT Press.