Learning from the Bible: Computational modeling of the costs of letter transpositions and letter exchanges in reading Classical Hebrew and Modern English

R. Harald Baayen
University of Tübingen & University of Alberta

Abstract

Letter transpositions are relatively harmless for reading English and other Indo-European languages with an alphabetic script, but severely disrupt comprehension in Hebrew. Furthermore, masked orthographic priming does not produce facilitation as in English (Frost, 2012). This simulation study compares the costs of letter transpositions and of letter exchanges for Modern English and Classical Hebrew, using the framework of naive discriminative learning (Baayen, Milin, Filipovic Durdjevic, Hendrix, & Marelli, 2011). The greater disruption of transpositions for Hebrew as compared to English is correctly replicated by the model, as is the relative immunity of loanwords in Hebrew to letter transpositions. Furthermore, the absence of facilitation of form priming in Hebrew is correctly predicted. The results confirm the hypothesis that the distributional statistics of the orthographic cues in the two languages are the crucial factor determining the experimental hallmarks of orthographic processing, as argued by Frost (2012).

In a recent study, Frost (2012) calls attention to an intriguing difference between Hebrew and English. English readers suffer little when letters are transposed or exchanged. For instance, *anwser* was found to provide the same amount of priming for the target *ANSWER* as the correct orthographic form, *answer*, itself (Forster, Davis, Schoknecht, & Carter, 1987; Perea & Lupker, 2003). Furthermore, a masked prime such as `mature` has been found to prime targets such as `NATURE`, subject to the condition that the words are from low-density similarity neighborhoods (Forster & Davis, 1991). By contrast, in Hebrew, letter exchanges

---

This research was supported by the Alexander von Humboldt foundation in the framework of an Alexander von Humbold research award, endowed by the Federal Ministry of Education and Research, to the author. The author is indebted to Ram Frost and Michael Ramscar for their comments on earlier versions of this study.

and letter transpositions are disruptive. For letter exchanges, no priming is observed, and letter transpositions (Frost, 2012; Velan & Frost, 2011) were detrimental to reading with a substantial drop in reading performance.

Frost (2012) argues that Hebrew requires a rigid orthographic coding scheme which contrasts with the flexible coding scheme that would be required for English or French. He also argues that the underlying neuro-circuitry of the visual system must operate on the same principles across both Semitic and non-Semitic languages. Frost's hypothesis is that it is differences in the statistical regularities across the two language families, Semitic and Indo-European, that give rise to different coding schemes, which themselves would be the result of the same underlying cognitive and neurobiological principles responding to very different language input.

The goal of the present study is to use computational modeling to examine Frost's hypothesis that the specific statistical properties of Hebrew are at issue. The method used to test the role of the distributional properties of Semitic and non-Semitic languages is corpus-driven computational modeling. The modeling framework adopted is that of Naive Discriminative Learning, henceforth NDL (Baayen et al., 2011; Baayen, 2011; Hendrix & Baayen, 2012). The NDL model implements a simple two-layer network with as orthographic input CUES letter unigrams and bigrams, and as output OUTCOMES lexical and grammatical meanings (e.g., 'hand', 'plural'). The model does not implement fuzzy encoding of orthography. That is, for a word such as *hand* the only letter bigrams taken into consideration are `#h, ha, an, nd, d#`. Furthermore, all letter unigrams and letter bigrams are assumed to become available at the same time.

Each OUTCOME is linked to all CUES. Hence, a subnet consisting of a given OUTCOME and its CUES is therefore formally a perceptron. The weights are not set by iteration over time, instead, for a given perceptron subnet the weights are determined by means of the equilibrium equations (Danks, 2003) of the Rescorla-Wagner equations (Wagner & Rescorla, 1972) for discriminative learning. As each perceptron is independent of all other perceptrons, the model instantiates *naive* discrimination learning, in the sense of naive Bayes classifiers.

Given corpus-derived matrices of co-occurrences of CUES and OUTCOMES, the sets of equilibrium equations can be solved straightforwardly using the Moore-Penrose pseudoinverse. Crucially, the weights are completely determined by the distributional properties of its input space. There are no free parameters, no parameters for decay rates, threshold activation levels, or resting activation levels. Baayen (2011) shows that naive discriminative learning, used as a classifier for data on the dative alternation, actually outperforms a generalized linear mixed model, and performs as well as a support vector machine.

The simulated latencies predicted by the naive discriminative reader (henceforth NDR) reflect a wide variety of distributional effects, including not only whole word frequency effects, but also morphological family size effects (Moscoso del Prado Martín, Bertram, Häikiö, Schreuder, & Baayen, 2004), inflectional entropy effects (Baayen, Feldman, & Schreuder, 2006), constituent frequency effects (Baayen, Kuperman, & Bertram, 2010), and paradigmatic entropy effects (Milin, Filipović Durdević, & Moscoso del Prado Martín, 2009). This is accomplished without any representations for complex words whatsoever — the model is a full semantic decomposition model. Furthermore, Hendrix and Baayen (2012) show that the model also predicts the n-gram frequency effects reported by Arnon and Snider (2010), again without the presence of representations for word n-grams.

The modeling results support the idea that morphological form representations mediating the flow of activation from orthographic units to semantics are unnecessary. (Baayen et al., 2011) discuss a great many arguments from theoretical linguistics in support of this claim. Hebrew presents a particularly strong test of this idea, because it is generally believed that the consonantal root, as well as the vocalic pattern, are representational units that must be accessed during reading for comprehension to take place (Frost, 2012). If the NDR model is on the right track, then it should be able to correctly predict the detrimental effects of letter exchanges and transpositions for Hebrew *without* the mediation of root and pattern morphemes. As hidden layers in connectionist models often function as implicit distributed representations mediating a mapping, our claim is that a hidden layer for representing roots and patterns is superfluous, in addition to being statistically dysfunctional.

Because the NDR is a tool for studying the relation between form and meaning, it is crucial to have semantically interpreted corpora available for training. For English corpora, a (highly simplified) semantic representation can be obtained by identifying inflectional functions, by isolating derivational meanings, and by taking the lemma as a pointer to a word's lexical meaning. In this way, a word such as *writers* receives the following set of meaning OUTCOMES: WRITE, AGENT, PLURAL. For English, tagged corpora and lexical databases facilitate the calculation of such meaning sets. For modern Hebrew, no such resources are available at present. Fortunately, there is a corpus of Classical Hebrew, namely, the collection of texts referred to as the Tenach in Judaism, and as the Old Testament in Christianity. These biblical texts are available on-line with literal morpheme-by-morpheme transliterations in English. By going back to the Bible, we obtain with these transliterations exactly the required semantic representations for setting the cue-to-outcome association strengths.

In what follows, we first describe the English and Classical Hebrew corpora in more detail, and then compare the distributions of unigrams and bigrams in these two languages. Next we proceed with an inspection of the weights and the predicted semantic activations. Against this background knowledge, we then consider the consequences of letter transpositions for Hebrew and English, as well as for non-Semitic loanwords in Hebrew. A simulation study of form priming with a single letter transposition is presented next, after which we conclude with a discussion of the results.

Training data sets for Hebrew and English

The simulation experiments reported in this study are based on two corpora, a corpus of classical Hebrew narrative texts, and for English the British National Corpus (Burnard, 1995).

Hebrew

From the interlinear Classical Hebrew texts at `http://www.scripture4all.org/OnlineInterlinear/Hebrew_Index.htm`, the older narrative texts (Genesis, Exodus, Judges, the two books of Samuel, the two books of Kings, Ruth, and Esther) and the legal texts from the Tenach (Leviticus, Numbers, Deuteronomy) were selected. The pdf files were converted into ASCII with the UNIX utility `pdftotext`, digraphs were converted into unigrams, and Hebrew words (defined as strings of letters bounded on either side by

braSiT	in-begin-ing
bra	he-create-ed
alHim	Elohim
aT	accusative marker
HSmim	the-heaven-s
uaT	and-accusative marker
HarZ	the-earth

Table 1: The first sentence from the book of Genesis: The Hebrew transliteration (left) and the English morpheme-by-morpheme glosses (right).

a space character) were aligned with their morpheme-by-morpheme English glosses. These English morphemic glosses will henceforth be used as the meaning components of a Hebrew word. Glosses for multimorphemic English translation equivalents were parsed into their constituent morphemes by look-up in the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995). Table 1 presents an example of the Hebrew-English alignments obtained. Due to inconsistencies in the mapping from pdf to ASCII format that resisted straightforward automated correction, the total number of alignments was reduced by roughly half to 113725.

Of the 113094 alignments, a minority (33538) had a single monomorphemic English translation equivalent. Due to cliticization, extensive person/number marking, and the realization of direct object and genitive pronouns as suffixes, a large majority of Hebrew strings required multi-word English glosses. Each of the glosses constitutes a meaning (or OUTCOME) for the naive discriminative learning model.

English

Count	Hebrew	English
1	33538	128
2	36487	4980
3	24081	38010
4	12886	47013
5	5047	15202
6	836	4402
7	182	2745
8	37	614

Table 2: Counts of number of meanings for the Hebrew words and English n-grams.

From the database of phrases from the British National Corpus compiled by Baayen et al. (2011), 113094 instances were selected at random, to provide an equal-sized instance base for English. This set of instances comprised 90780 word trigrams and 22314 word

Form	Orthography	Root	Root bigrams
lamad	`lmd`	`lmd`	`lm, md`
lamadti	`lmdti`	`lmd`	`lm, md`
lmadtem	`lmdtm`	`lmd`	`lm, md`
yilmad	`ilmd`	`lmd`	`lm, md`
limmed	`lmd`	`lmd`	`lm, md`
talmud	`tlmwd`	`lmd`	`lm`

Table 3: Examples of Classical Hebrew forms of the root *lmd* ('to learn'), their orthography, and bigrams that are part of the root (*lm, md*).

bigrams. In this data set, a word can be morphologically complex, consisting at most of two monomorphemic stems and an inflectional affix. The choice of word bigrams and trigrams for English is motivated by the consideration that most Hebrew letter strings have multi-word translation equivalents in English. For instance, Hebrew *wydbr* translates into *and he said*. By estimating weights from word n-grams rather than from single words, a learning bias favoring English is avoided. It should be noted, however, that the two corpora are not well balanced with respect to the distribution of the numbers of meanings for a given string, as shown in Table 2. For typologically very different languages such as Hebrew and English, such imbalances are inevitable.

Comparison of distributions of unigrams and bigrams

Frost (2012) describes the orthographic coding scheme of Hebrew print as relying "mainly on the few letters that carry morphological information, whereas the other letters of the word do not serve for lexical access, at least not initially" (p. 21). He further argues that for Hebrew, but not for Indo-European languages, individual letters of base words carry semantic values (p. 23). In the light of these claims, it is worth considering the distributions of Hebrew and English unigrams and bigrams.

A comparison (ignoring the space character) of the distribution of letter unigram frequencies for Hebrew and English reveals two very similar distributions that do not differ significantly according to a Kolmogorov-Smirnov test (p =0.5544). This finding argues against the idea in Hebrew, unlike in English, some letters would be especially informative. It is also not the case that there are letters that are not used as part of a root. Some letters, such as *w, i* and *h*, when used as root consonants, may not be present in some of the written forms of a given (weak) verb's inflectional paradigm, but they will be visible in at least some forms.

Letter unigrams by themselves, therefore, are not likely to offer an explanation for the different processing characteristics of Hebrew and English. Letter bigrams, by contrast, are more promising, as bigrams capture important chunks of a word's (mostly tri-consonantal) root. As can be seen in Table 3, across very different spoken forms, thanks to the omission of many vowels in the orthography, at least one root bigram will be present in most forms.

If root bigrams play a role as cues to root meanings, then we may expect that the distributions of Hebrew and English bigrams do differ significantly. More specifically, due

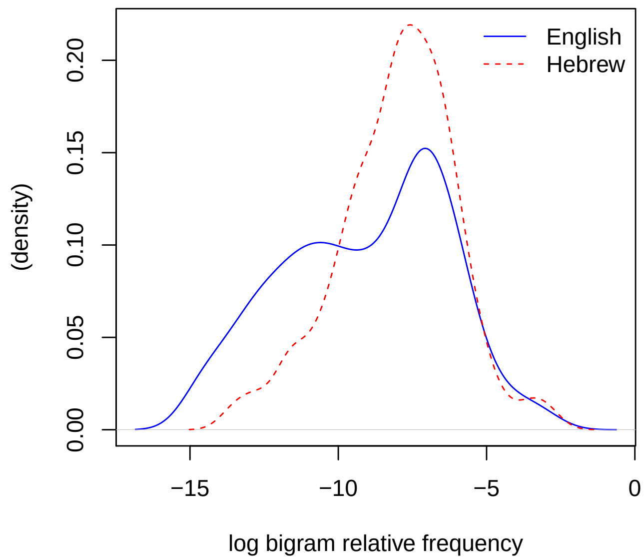

Figure 1. Densities of letter bigram frequencies for Hebrew (solid) and English (dashed).

to the recurrence of the same bigram across the many different derivational and inflectional forms of a root, Hebrew should reveal a distribution of bigrams with more higher-frequency bigrams compared to English.

This is exactly what Figure 1 reveals: There is a highly significant difference between the bigram distributions of these two languages ($p = 0$, Kolmogorov-Smirnov test), such that higher bigram frequencies are more characteristic of Hebrew ($p < 0.0001$, Wilcoxon test). Furthermore, since Hebrew has fewer distinct bigrams (556) as compared to English (650), Hebrew can be said to use its smaller set of bigrams more intensively. This result is fits well with Frost's claim (2012, p.27) that

> for efficient reading, the statistical properties of letter distributions of the language, and their relative contribution to meaning, have to be picked up, and the transitional probabilities of letter sequences have to be implicitly assimilated.

Although bigram frequencies represent transitional probabilities, by themselves they are completely blind to the link between orthographic form and a root's meaning. The naive discriminative learning framework offers the possibility to assess the strengths with which bigrams (and also unigrams) are linked to (root) meanings, by means of the cue-to-outcome association weights on the links from orthographic cues to the meaning outcomes.

Orthography-to-meaning association weights

The `ndl` package (Arppe, Milin, Hendrix, & Baayen, 2011) was used to set the weights from the orthographic cues (letter unigrams and bigrams) of Hebrew words and English word n-grams to their component meanings.

For each language, the cues consisted of letter unigrams and bigrams. For Hebrew *HarZ*, the set of cues contained the unigrams *H, a, r, Z* and the pairwise bigrams *#H*, *Ha*, *ar*, *rZ*, and *Z#*. The English three-word sequence *and he said* was coded as `#and#he#said#`, and yielded the cue set {*#, a, n, d, h, e, s, i, #a, nd, d#, #h, he, e#, #s, sa, id, d#*}. The cue-to-outcome association weights were estimated using the equilibrium equations for the Rescorla-Wagner equations of Danks (2003).

Language	Gram	Mean	Median	Count
Hebrew	unigram	0.000013	-0.0000099	92483
Hebrew	bigram	0.000088	0.0000040	2143193
English	unigram	0.000013	-0.0000015	173623
English	bigram	0.000028	0.0000000	3717927

Table 4: Statistics for unigram and bigram association weights for Hebrew and English

Table 4 presents mean and median of the unigram weights and bigram weights for Hebrew and English. The median unigram weights are negative for both languages, and smaller than the bigram weights. The unigram weights are uninformative about lexical meanings. Too many meanings link up to the same unigrams for the unigrams to be useful as discriminative cues.

Interestingly, a striking difference emerges between Hebrew and English. For Hebrew, the median weight for bigrams is half of the median for the unigrams, but with the sign reversed. By contrast, the English bigram weights are positive, but close to zero. This suggests that the functional load for Hebrew bigrams is much greater.

It is worth noting that the median number of letters in a Hebrew word is 4, whereas the median number of letters in an English n-gram, not counting spaces, is 13. As a consequence, a larger number of cue-to-outcome connections will contribute to a meaning activation for English compared to Hebrew. Furthermore, experience with modeling English reading indicates that adding bigram cues to unigram cues leads to enhanced prediction accuracy: They are not superfluous. Yet even if we correct for two orders of magnitude to take word length differences into account, the magnitude of Hebrew bigram cues remains many orders of magnitude greater compared to English cues.

There is also a difference in the number of meanings that are distinguished in the training data for the two languages, and it might be argued that this difference is at issue. However, the magnitude of this difference is moderate: 4021 for Hebrew and 5987 for English: The Hebrew data set has approximately two thirds of the number of meanings distinguished in the English data set. More importantly, as the NDL model implements *naive* discriminative learning, for which each meaning and its cues are evaluated independently of all other meanings, the difference in the number of meanings cannot explain the differences in weights.

These considerations taken together suggest that for Hebrew the functional load of bigram-to-meaning weights is likely to be much higher compared to English, exactly as predicted by Frost (2012).

Meaning activations

The weight matrices produced by the NDR specify the assocation strengths of unigram and bigram cues to outcomes (meanings), allowing the activations of these meanings given an input string to be straightforwardly estimated by summation of the incoming weights from active cues. For Hebrew, activations were evaluated for 700 verbs in the Hebrew corpus, as verbs show the greatest inflectional variation and hence provide the best and most difficult test case for naive discriminative learning. (Verb status was determined on the basis of the English gloss containing a lexeme that according to CELEX can be a verb, followed by manual pruning of, for instance, unlikely conversion verbs.) For English, 700 content words were selected, which were matched with the Hebrew words in length (p =0.9994, Kolmogorov-Smirnov test). The English (lemmatized) words, however, were more frequent ($p < 0.0001$, Wilcoxon test; median Hebrew frequency: 5; median English frequency: 16).

	Hebrew	English
median activation	0.0328	0.0124
median activation from unigrams	0.0013	0.0001
median activation from bigrams	0.0274	0.0122
median activation unigrams	0.0011	0.0003

Table 5: Median activation, median activation from unigram cues, median activation from bigram cues, and the median activation from a model with unigram cues only, for Hebrew and English.

The first row of Table 5 shows that, as expected given the greater magnitude of Hebrew weights, the median activation for Hebrew is greater than the median activation for English. Note that as the length distributions of the Hebrew and English test words do not differ significantly, it is unlikely that Hebrew activations would be higher due to Hebrew words containing more bigrams. Furthermore, the frequencies of the Hebrew words were lower than those of the English words, hence frequency cannot explain the observed pattern of results as frequency by itself predicts greater activation for English, and smaller activation for Hebrew, contrary to fact.

For both Hebrew and English, the bulk of the activation is contributed by the active bigram cues — the contribution of unigram cues is much smaller, as indicated by rows two and three of Table 5. The last row of this table lists the activations that are obtained when a model is trained with only unigram cues and no bigram cues whatsoever. Such a unigram-only model gives rise to reduced activations similar to those observed in the unigram+bigram model from just the unigram cues.

It is clear from Table 5 that the bigram cues are the key contributors to the meaning activations. For both languages, the order information captured by the bigrams is a key factor in the mapping of form to meaning. The set of bigram cues is large enough so that individual bigrams can become good cues for individual meanings. At the same time, this set is not too large to become computationally prohibitive.

In summary, Hebrew uses a smaller set of bigram cues more intensively than English, and these bigram cues tend to have stronger associative weights to lexical meanings compared to English. This gives rise to higher activation levels for these meanings, even where both sets of words are of similar length, and even though Hebrew words (more specifically, their meanings) are characterized by lower frequencies. In what follows, we address whether these remarkable properties of the Hebrew bigrams can explain the very different priming and letter transposition effects observed for Hebrew and English.

Letter transpositions

How does a letter transposition affect the activation of a word's meaning in Hebrew compared to English? To address this question, the activations of the Hebrew verbs were estimated for when a single letter pair was transposed, subject to the condition that the letter pair was not word initial nor word final. This restriction limited the set of eligible verbs to the 588 verbs with a minimal length of four letters (including clitics and affixes). For English, 666 words, also with a word length exceeding three characters, were subjected to the transposition of a non-edge-aligned letter pair. For both languages, the respective cue-to-outcome weight matrices were used to calculate the activation of the words' meanings, both for the standard form, and for the form with a letter transposition.

	Hebrew	English
median activation	0.0344	0.0124
median activation with transposition	0.0026	0.0035
median difference	0.0238	0.0036
proportional reduction in activation	0.9237	0.7186

Table 6: Median activation, median activation with a single transposition of one non-edge aligned letter pair, the median of the difference between the two, and the proportional reduction in median activation, for Hebrew and English. The number of observations is 588 for Hebrew and 642 for English.

Disruption by letter transpositions can be gauged by calculating the extent to which a word's activation decreases due to a letter transposition. Table 6 clarifies that the amount of disruption caused by a single letter transposition is indeed greater for Hebrew than for English: The median disruption in Hebrew (0.0238) is 6.6 times that for English (0.0036, p <0.0001, Wilcoxon test on the proportional reductions).

	Estimate	Std. Error	t value
Intercept (Language=English, Treatment=None)	904.42	2.86	316.36
Language = Hebrew	-15.96	4.16	-3.84
Treatment = Transposition	20.23	1.96	10.30
Language = Hebrew & Type = Transposition	23.91	2.87	8.33

Table 7: Coefficients of a mixed-effects regression model for the simulated reading times of Hebrew and English words with and without a letter transposition.

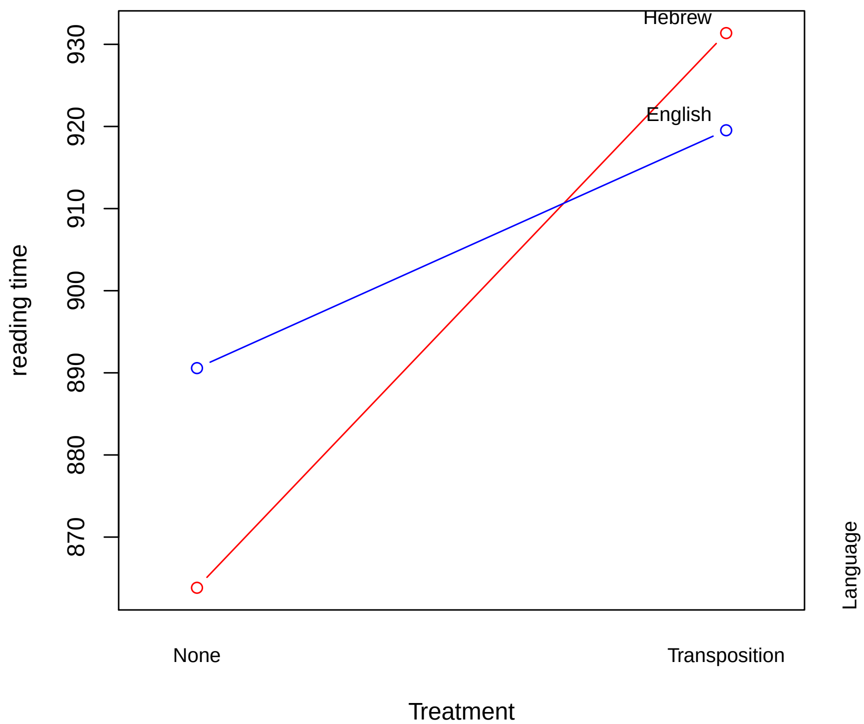

Figure 2. Mean simulated reading times for Hebrew and English with and without a letter transposition, as estimated by a linear mixed-effects model with random intercepts for Item.

Chronometric methods assess meaning activations indirectly through response variables such as response latencies and reading times. To simulate this, we define reading times as the log of the reciprocal of the activation. To do so, it is necessary to avoid negative activations, which can be achieved by addition of the absolute value of the minimum activation. It is also necessary to back of from zero, which we do by adding one:

$$\mathrm{t} = \log \frac{1}{a + \mathrm{I}_{[\min(a) \leq 0]} |\min(a)| + 1}. \tag{1}$$

For ease of interpretation, these reading times are subsequently rescaled to the interval [200,1000] ms.

A mixed-effects analysis of the simulated reaction times with Item as random-effect factor was fitted to the data. Subsequently, outliers with absolute standardized residuals exceeding 2.5 were removed and the model refitted. The trimmed model, which is summarized in Table 7 and visualized in Figure 2, revealed a significant interaction of Language by Treatment, with a transposition cost for Hebrew that is twice that for English. The results of this simulation therefore fit with the observed much greater difficulty that Hebrew readers experience with words with transposed letters.

Loanwords in Hebrew

Hebrew transliteration	English source	Hebrew transliteration	English source
trus	truss	tSaild	child
mudrn	modern	labradur	labrador
braS	brash	kuuta	quota
asbstus	asbestos	abril	april
spak	speck	uarrn	warren
brut	brute	miriad	myriad
rumans	romance	tannin	tannin
suap	sweep	starb	starve
sulu	solve	skrabbl	scrabble
fauntn	fountain	hurd	idiom
blutS	blotch	kurdl	curdle
suldr	solder	kambur	camphor
skrambl	scramble	traS	trash
Hurd	hoard	skil	skill
kautS	couch	Hurd	horde
glatin	gelatine	sulbunamid	sulphonamide
plank	plank	sugst	suggest
snaipr	sniper	dmagug	demagogue
bras	brass	gardn	garden
pulip	polyp	buru	bureau

Table 8: English sources and Hebrew transliteration of 40 randomly selected loanwords

Loanwords in Hebrew are reported not to suffer from adverse effects of letter transpositions (Velan & Frost, 2011; Frost, 2012). To study whether the naive discriminative learning framework helps explain the remarkable immunity of loanwords against transpositions, 40 English content words were selected randomly from the English dataset, subject to the condition that their meanings do not occur in the Hebrew dataset. For each of these words, a transliteration into the Hebrew orthography was made (see Table 8). The loanwords were added to the Hebrew dataset, and were assigned frequencies randomly selected from the frequency list of meanings in the outcomes vector of the Hebrew dataset. As no context is available in the original corpus, no further context was added to the loanwords. Subsequently, the weights and activation matrices were recalculated.

	Normal	Transposed
Hebrew	0.0344	0.0027
Loanwords	0.1975	0.0547

Table 9: Median activation for native Hebrew words and 40 loanwords, without and with transposition.

Table 9 shows that the median activation for loanwords, 0.197 is substantially larger

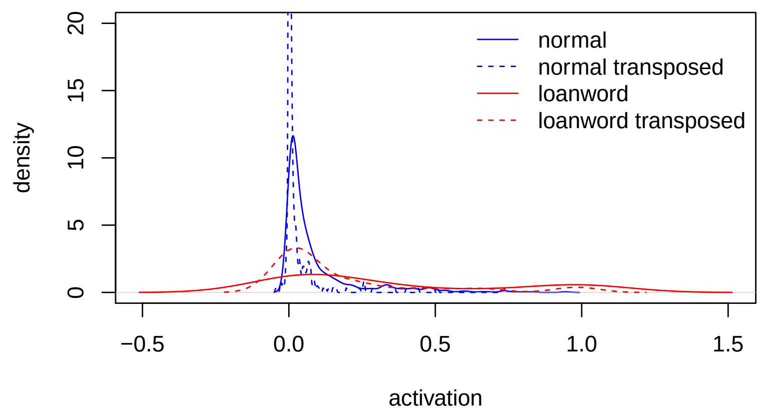

Figure 3. Estimated densities for normal (black) and loan words (grey) with (dashed line) and without (solid line) transposition.

than the median activation for native words, 0.034 ($p = 0.0001$, Wilcoxon test).

Letter transpositions reduce the median activation for the loanwords from 0.1975 to 0.0547. Crucially, the median activation for loanwords after a letter transposition, 0.0547 is as high as that of non-loans without transposition (0.0344, a non-significant difference according to a Wilcoxon test).

Figure 3 presents the estimated densities for non-loans and loanwords with and without transpositions. The density for tranposed loanwords is quite close to that of the untransposed nonwords, in conformity with the medians in Table 9.

	Estimate	Std. Error	t value
Intercept (Type = Native, Treatment = None)	885.5293	3.5285	250.9661
Type = Loanword	-142.8045	13.4210	-10.6404
Treatment = Transposed	47.4346	2.6931	17.6131

Table 10: Coefficients of a linear mixed-effects model fit to the simulated RTs for normal and loan words, with and without transposition.

A mixed-effects model fitted to the simulated RTs, obtained in the same way as above using (1) and rescaled to the $[200, 1000]$ ms interval, is summarized in Table 10 and Figure 4. (In order to avoid violating normality assumptions, outliers with absolute standardized residuals were removed from the data set and the model refitted. As a consequence, a minor interaction of Transposition by Loanword status lost significance.) According to this trimmed model, transposition has a main effect (a cost of 47 ms) and loanwords have, on average, an advantage of 119 ms. Even after trimming, the distribution of RTs remained

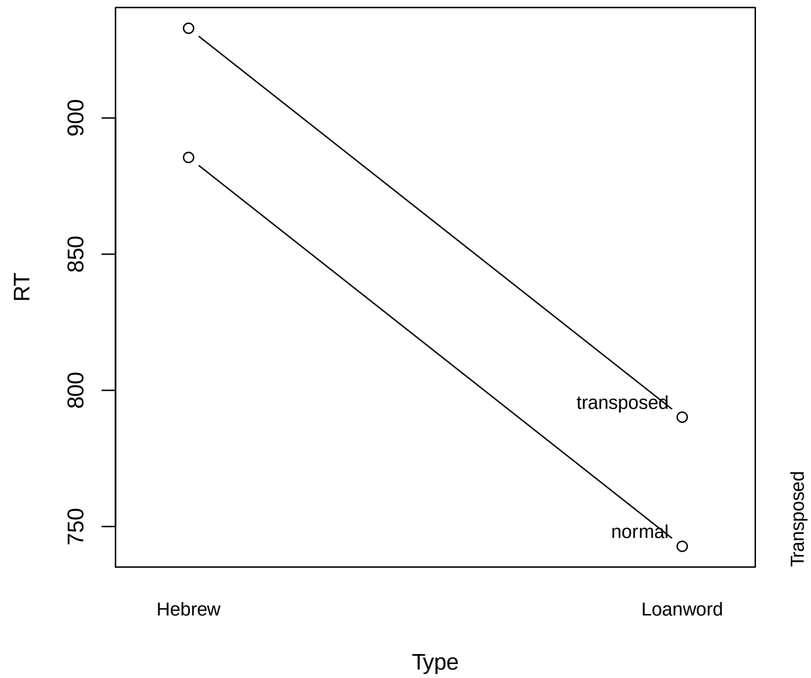

Figure 4. Estimated simulated mean RTs for Hebrew words and loanwords with and without transposition.

very skewed, due to a very skewed — but ecologically valid — word frequency distribution.

These results indicate that the model succesfully captures the relative immunity of loanwords to letter transpositions. Interestingly, if the present model is on the right track, the cause of this immunity is not that loanwords would not suffer from transpositions. To the contrary, as can be seen in Figure 4, they suffer exactly to the same degree as native words. The reason that transposed loans appear immune is that loanwords as such are easier to read. This reading advantage masks the costs of the transpositions. The reason for the reading advantage for loanwords is easy to trace. Loanwords have unusual letter combinations. Unusual bigrams are excellent cues for the loanword meanings, which are learned better, and therefore receive higher activations and shorter reading times. Note that the model predicts that untransposed loanwords can be read faster than untransposed native words, other things being equal. Whether this prediction is correct awaits further empirical verification with real loanwords rather than the constructed loanwords used in the present simulation study.

At present, however, our results suggest that the difference between loanwords and native words resides primarily in the fact that the writing system for native words has removed nearly all redundancy, leaving a system in which the cues that have been left carry a high functional load. With redundancy removed, the system is not robust to noise in the signal, as witnessed by the extreme disruption caused by letter transpositions. Because loanwords do not have the Semitic phonotactics, redundancy is re-introduced, and this in turn allows for faster and more robust processing.

Simulating priming

The preceding section clarified that the different effects of letter transpositions across Hebrew and English can be understood perfectly well within the naive discriminative learning approach. This section addresses whether this framework makes the correct predictions for the effects of form priming. The central empirical finding is that form related primes, obtained for instance by random replacement of a letter, produce facilitation compared to an unrelated baseline in English, but not in Hebrew.

For the evaluation of priming effects, the same 700 English and Hebrew words were considered as for the transposition simulation experiment. Two conditions were considered in addition to the identity condition, a priming condition in which the prime was obtained from a target word with a random substitution of one letter, and an unrelated condition in which the prime was a randomly selected word of the same length. As for English the longest length (12) was represented by a single word, no replacement was available. As a consequence, the number of English words studied was reduced by one.

Priming was modeled, as in Baayen et al. (2011), with the compound cue strength model CCS of Ratcliff and McKoon (1988), according to which the joint cue strength of prime P and target T is given by the sum of all relevant activations for prime a_{P_i} and target a_{T_i} given a prime weight w representing the duration of the prime:

$$\text{CCS} = \sum_i a_{P_i}^{(w)} a_{T_i}^{(1-w)} \quad (0 \leq w \leq 0.5). \tag{2}$$

For the identity condition, the prime is the target, and the identity condition CSSi reduces to the activation of the target a_T. For the prime condition, in which the letter exchange results in a nonword, the relevant terms are a_{PT}, the activation of the target meaning given the prime, and a_{TT}, the activation of the target meaning given the target. The compound cue strength for the letter-exchange prime and its target, CSS_e, is

$$\text{CCS}_e = a_{PT}^{(w)} a_{TT}^{(1-w)} \quad (0 \leq w \leq 0.5). \tag{3}$$

For the unrelated condition, there are two sources of confusion. The meaning of the unrelated prime can interfere with the meaning of the target, and the unrelated prime form can interfere with the activation of the target meaning. Both interferences were modeled with (2), with the index of the sum ranging over both sources of confusion. Let a_{PP} denote the activation of the prime meaning by the prime. The compound cue strength for the unrelated condition, CCS_u can now be defined as

$$\begin{aligned} \text{CSS}_u &= a_{PP}^{(w)} a_{TT}^{(1-w)} + a_{PT}^{(w)} a_{TT}^{(1-w)} \\ &= [a_{PP}^{(w)} + a_{PT}^{(w)}] a_{TT}^{(1-w)}, \end{aligned} \tag{4}$$

subject, as before, to the condition $(0 \leq w \leq 0.5)$. These equations are defined only for positive a_i. For $a < 0$, a^w was defined as $-(|a|^w)$. The activations for the prime condition,

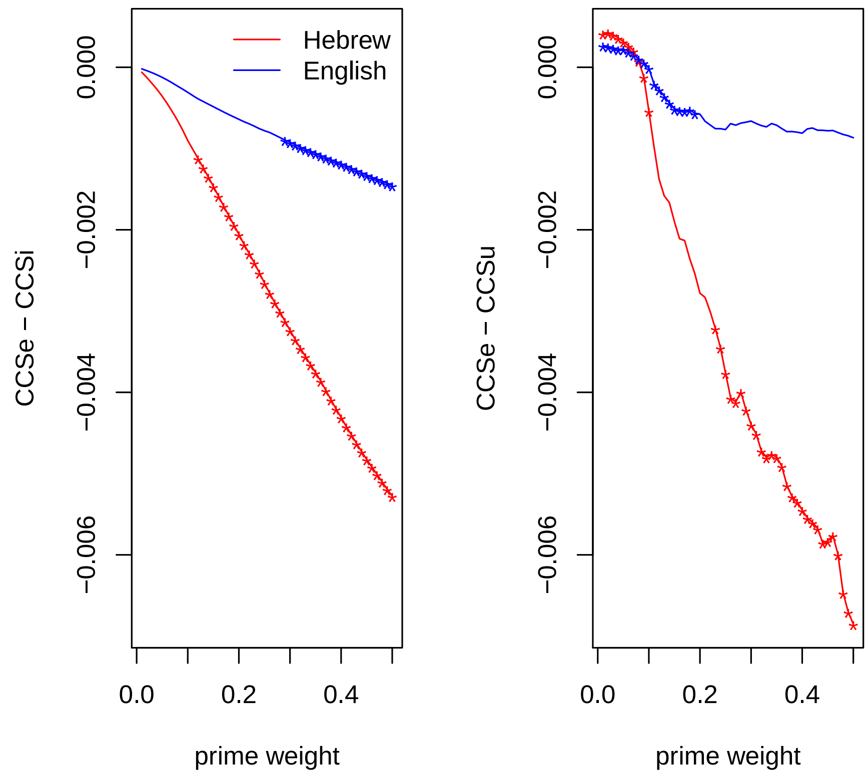

Figure 5. Difference in the compound cue strength of a single letter exchange prime CCSe and the compound cue strength for the identity condition CCSi (left) and the compound cue strength of the unrelated condition CCSu (right) for Hebrew (red) and English (blue). Points indicate significance according to a Wilcoxon test.

the identity condition, and the unrelated condition were rescaled to the [0,1] interval, for Hebrew and for English separately.

Figure 5 presents, for the full range of possible values of the prime weight w, the difference in the compound cue strength of a one-letter exchange prime CCS_e and the compound cue strength for the identity prime CCS_i (left) and the compound cue strength for the unrelated prime CCS_u (right), evaluated in the median. The left panel clarifies that a letter replacement is substantially more detrimental for Hebrew than for English. A non-identical prime blurs the signal (Norris & Kinoshita, 2008), and for Hebrew this blurring is substantially more harmful than for English. Not only is the magnitude of the effect larger for Hebrew, but the effect reaches significance for smaller w, that is, for shorter prime durations.

The right panel of Figure 5 indicates that for very small prime durations w the compound cue strength of the letter exchange prime is greater than that of an unrelated prime. This effect is significant in both English and Hebrew, but it is stronger for Hebrew. For increasing prime duration (w), the effect reverses: The unrelated prime (which is a word) has a stronger compound cue strength than the letter-exchange nonword prime.

	Estimate	Std. Error	t value
Intercept (Language=Hebrew, Type=Unrelated	813.1404	4.0901	198.8053
Language=English	33.7517	5.7629	5.8567
Type=Identity	-7.4045	1.4443	-5.1267
Type=Letter Exchange	4.4040	1.4460	3.0458
Language=English & Type=Identity	-0.7934	2.0368	-0.3895
Language=English & Type=Letter Exchange	-7.9405	2.0356	-3.9008

Table 11: Simulated RTs for the priming data. The reference level for Language is Hebrew, and the reference level for Type is Unrelated

As before, we take the log of the reciprocals of the simulated activations to obtain masked priming response latencies. Figure 6 and Table 11 summarize the results of a linear mixed-effects model with Item as random-effect factor fitted to these simulated reaction times with prime weight $w = 0.15$. (Eight extremely small activations were reset to 0.15) followed by rescaling to the interval [200,1000]. After a first fit of the model, potentially harmful outliers were identified and removed, and the model was refitted.

The model indicates that the identity condition elicits shorter simulated reaction times than the unrelated condition for both languages, without an interaction. For English, the letter-exchange prime elicits significantly shorter reaction times than the unrelated prime. For Hebrew, by contrast, small but significant inhibition emerges. This inhibitory effect can be rendered non-significant by adding Gaussian response noise with a standard deviation of 50, without affecting the significance of the facilitation for English primes compared to the unrelated condition.

Root priming

One important question that remains is whether naive discriminative learning can also account for root priming and pattern priming. Root priming concerns prime-target pairs which share the same root, e.g. for the root `lmd` the forms `lmdti` (first persion perfective) and `azmr` (first person imperfective). Pattern priming involves forms with different roots that share the same vowel+affix pattern, for instance, `iSmr` (*yishmor*) and `ilmd` (*yilmad*). To address this issue, first consider the predictions of the compound cue strength model of Ratcliff and McKoon (1988) under the assumption that prime and target pairs are balanced with respect to their distributional properties. For balanced primes and targets, it can be shown that (2) leads straightforwardly to the compound cue strengths listed in Table 12. These activations correctly predict the longest response latencies for the unrelated condition, intermediate latencies for the root and pattern priming conditions, and the shortest latencies for the identity condition. (In reality, perfect matching of primes and targets is impossible, and the expected values in Table 12 will vary substantially depending on a words position in

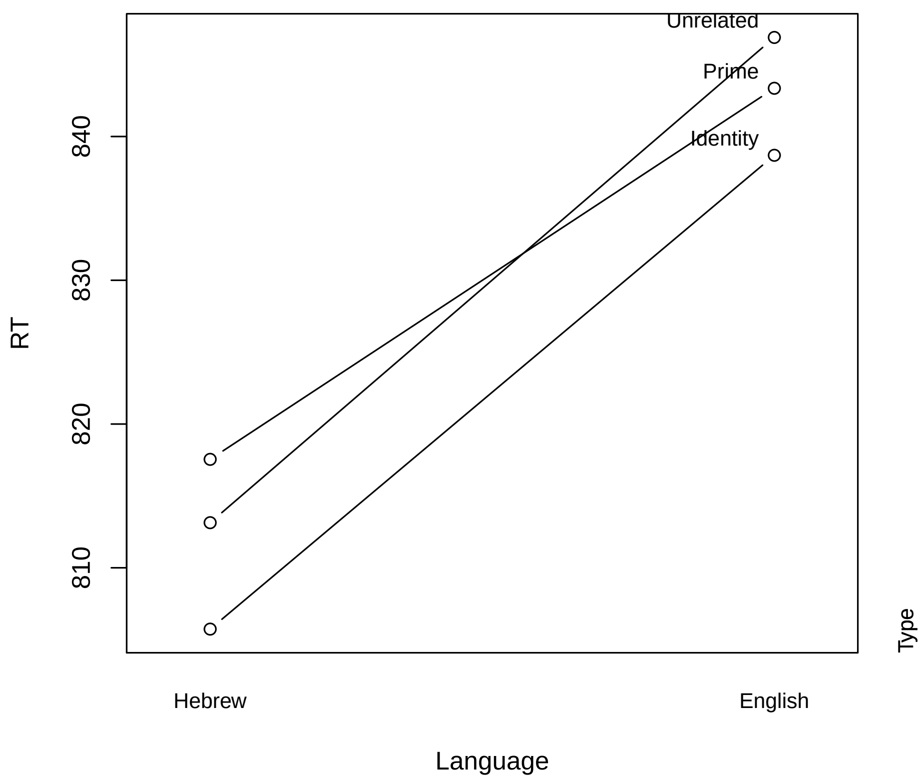

Figure 6. Estimated means for simulated reading latencies for Hebrew and English crossed with prime condition (Identity, Prime (letter exchange), Unrelated).

its distributional hyperspace. Actual computational modeling will be essential for clarifying whether naive discriminative learning will indeed be succesful in modeling pattern priming.)

For the identity condition, the compound cue strength is predicted to be the sum of the target's lexical and grammatical meanings. For the unrelated condition, the compound cue strength is very small, and close to zero.

For root priming, the compound cue strength is a proportion of the target's lexical activation. This proportion will depend primarily on the number of shared root bigrams and their cue strengths.

Pattern priming gives rise to a compound cue strength that is a proportion of the activation of the target's grammatical meaning. The magnitude of this proportion probably depends less on letter bigrams than is the case for roots, for two reasons. First, patterns are often co-indicated in the orthography by affixes, compare, for instance, the forms `ilmd` (*yilmad*) and `iSmr` (*yishmor*), both third person singular imperfective forms. The patterns *yi-o* and *yi-a* are carried solely by the word-initial and invariant `i` in the orthography. For this initial `i`, the edge bigram `#i` is a potential cue, but due to preceding clitics, the `i` is often

Condition	Compound Cue Strength
identity	full activation lexical meaning + full activation grammatical meaning
root priming	partial activation lexical meaning
pattern priming	partial activation grammatical meaning
unrelated	near zero

Table 12: Compound cue strengths for four priming conditions under balanced prime-target pairs.

not word-initial, reducing the importance of the bigram, and enhancing the cue validity of the `i` unigram. Second, some patterns are carried exclusively by a vowel, as, for instance, in the case of the `w` in the absolute infinitive of Classical Hebrew (`lmwd`, *lamod*). Here, the unigram w is likely to be a relatively important cue. As a consequence, the processing of Hebrew patterns is expected to be less dependent on bigrams, and to be more similar to the processing of vocalic alternations in Germanic languages.

A natural extension that might be necessary is to include not only order relations at distance 1 (captured by bigrams), but also order relations at distance 2 (i.e., discontinuous bigrams such as `l-d`) and possibly even order relations at greater distances. Such extensions may be necessary not only for fine-tuning predictions for Semitic languages (the `w-#` discontinuous bigram would capture the pattern for the absolute infinitive, for instance), but may prove essential for properly modeling reduplication in languages such as Tagalog. It is important to note that such discontinuous bigrams maintain positional information, unlike "open bigram units" (Grainger & Whitney, 2004) which only encode relative position but provide no information about distance.

Whether such extensions are truly necessary awaits further modeling, especially because the beauty of naive discriminative learning and also of string kernels in machine learning is that these models do not take orthographic cues into account just by themselves, but in their co-occurrence configurations. For NDL, this implies that next to isolated unigrams and isolated bigrams, co-occurrence pairs — unigram-unigram, unigram-bigram, and bigram-bigram — are implicitly also considered. As a consequence, a word pattern such as exemplified by the Hebrew words *tizmoret, tiksoret, tisroket, tifzoret, tikrovet, tirkovet* will be detectable as long as the bigrams `ti` and `et` and the unigram `o` co-occur as cues for a consistent meaning. For Hebrew, we therefore anticipate that the full range of lexical processing effects can be understood with the present very basic cue set.

Discussion

The results of this simulation study provide the following insights: First, bigram cues have a greater functional load in Hebrew than in English. There are relatively more frequent bigrams in Hebrew, and the bigram cue-to-outcome association strengths in the NDL model are stronger in the model for Hebrew than in the model for English.

Second, letter transpositions emerged as more disruptive for Hebrew than for English. Furthermore, loanwords from English into Hebrew, adjusted to the Hebrew orthography, and assigned frequencies sampled from observed Hebrew frequencies, are equally disrupted by letter transpositions as non-loaned Hebrew words.

Third, near-identity orthographic priming generates facilitation (compared to an unrelated baseline) for English but a small inhibitory effect in Hebrew.

Fourth, the empirical findings can be approximated without positing lexical access is mediated by hierarchically organized intervening layers of syllable units and morpheme units (Dell, 1986; Taft, 1994; Schreuder & Baayen, 1995; Baayen, Dijkstra, & Schreuder, 1997). This results is especially striking because the massive array of priming studies for Hebrew and Arabic have been interpreted as supporting separate representations for consonantal root morphemes as well as for vocalic pattern morphemes (see, for instance, for Hebrew Bentin and Frost (1995); Frost, Forster, and Deutsch (1997); Frost, Deutsch, and Forster (2000); Frost, Kugler, Deutsch, and Forster (2005); Frost (2012) and for Arabic, e.g., Boudelaa and Marslen-Wilson (2001)).

Interestingly, Frost (2012) accepts as given that indeed root and pattern morphemes indeed mediate lexical access in Semitic. This forces him to assume that the flexible letter position coding that works so well within interactive activation frameworks for explaining priming and transposition effects in English must be language-specific. He therefore argues that for Hebrew and Arabic, some form of a rigid, position specific coding system must be in place. Without such a rigid coding system, it would be impossible to access the proper root and pattern morphemes, which in turn would be the gatekeepers to meaning.

In the naive discriminative learning approach pursued in the present study, which is basically parameter free (for priming, a parameter for the prime duration is required), there is no strict positional encoding. In fact, the problem of letter-position flexibility does not arise at all. Letter unigrams encode which letters are present, and letter bigrams encode relative positions of letter pairs. Relative position information turns out to carry a much heavier functional load in Hebrew than in English. The greater functional load of letter pairs translates straightforwardly into catastrophic effects of letter transpositions, and the absence of one-letter-exchange form priming.

It is therefore not necessary to argue that some languages would encode letter positions flexibly and others rigidly. Instead, all that is necessary is a sufficiently simple and open universal learning algorithm that is sensitive to the (potentially very different) distributional properties characterizing the mapping of form onto meaning in different languages. The crucial insight is that the similarity space provided by letter bigrams is sufficiently sparse to allow small sets of bigrams to become sufficiently reliable cues to jointly activate the intended meanings.

References

Arnon, I., & Snider, N. (2010). Syntactic probabilities affect pronunciation variation in spontaneous speech. *Journal of Memory and Language*, *62*, 67–82.

Arppe, A., Milin, P., Hendrix, P., & Baayen, R. H. (2011). ndl: Naive discriminative learning [Computer software manual]. Available from `http://CRAN.R-project.org/package=ndl` (R package version 0.1.1)

Baayen, R. H. (2011). Corpus linguistics and naive discriminative learning. *Brazilian Journal of Applied Linguistics*, *in press*.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, *36*, 94–117.

Baayen, R. H., Feldman, L., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, *53*, 496–512.

Baayen, R. H., Kuperman, V., & Bertram, R. (2010). Frequency effects in compound processing. In S. Scalise & I. Vogel (Eds.), *Compounding.* Amsterdam/Philadelphia: Benjamins.

Baayen, R. H., Milin, P., Filipovic Durdjevic, D., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*(3), 438-481.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (cd-rom).* University of Pennsylvania, Philadelphia, PA: Linguistic Data Consortium.

Bentin, S., & Frost, R. (1995). Morphological factors in visual word identification in Hebrew. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (p. 271-292). Hillsdale, NJ: Erlbaum.

Boudelaa, S., & Marslen-Wilson, W. D. (2001). Morphological units in the Arabic mental lexicon. *Cognition*, *81*(1), 65-92.

Burnard, L. (1995). *Users guide for the British National Corpus.* Oxford university computing service: British National Corpus consortium.

Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology*, *47*(2), 109–121.

Dell, G. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, *93*, 283-321.

Forster, K., & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. *Journal of Memory and Language*, *30*, 1–25.

Forster, K., Davis, C., Schoknecht, C., & Carter, R. (1987). Maskd priming with graphemically related forms: Repetition or partial activation. *The Quarterly Journal of Experimental Psychology*, *39*, 211–251.

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, in press.

Frost, R., Deutsch, A., & Forster, K. (2000). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 751-765.

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 829-856.

Frost, R., Kugler, T., Deutsch, A., & Forster, K. (2005). Orthographic Structure Versus Morphological Structure: Principles of Lexical Organization in a Given Language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1293–1326.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid read wrods as a wlohe? *Trends in Cognitive Sciences*, *8*, 58–59.

Hendrix, M., P.and Ramscar, & Baayen, R. H. (2012). Sidestepping the combinatorial explosion: Towards a processing model based on discriminative learning. *Language and Speech*, under revision.

Milin, P., Filipović Durdević, D., & Moscoso del Prado Martín, F. (2009). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language*, 50–64.

Moscoso del Prado Martín, F., Bertram, R., Häikiö, T., Schreuder, R., & Baayen, R. H. (2004). Morphological family size in a morphologically rich language: The case of Finnish compared to Dutch and Hebrew. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 1271–1278.

Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and bayesian inference: Insights from masked priming. *Journal of Experimental Psychology*, *137*(3), 434–455.

Perea, M., & Lupker, S. (2003). Does judge activate court? Ttransposed-letter similarity effects in masked associative priming. *Journal of Experimental Psychology: General*, *137*, 434–455.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*(3), 385–408.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing* (p. 131-154). Hillsdale, New Jersey: Lawrence Erlbaum.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, *9*(3), 271-294.

Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the natre of orthographic processing. *Cognition*, *118*, 141–156.

Wagner, A., & Rescorla, R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning ii* (pp. 64–99). New York: Appleton-Century-Crofts.