UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Syllable Priming and Lexical Representations: Evidence from Experiments and Simulations

Permalink
https://escholarship.org/uc/item/9hq4q0gm

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 14(0)

Author
Corina, David P.

Publication Date
1992

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Syllable Priming and Lexical Representations: Evidence from Experiments and Simulations

David P. Corina Ph.D.
U.S.C. Program in Neuroscience
HNB 18C, University Park, Los Angeles CA 90089-2520
corina@gizmo.usc.edu

Abstract.

This paper explores the composition of syllable structure in lexical representations. Data from auditory lexical decision experiments are presented which demonstrate that syllable structure is represented in the mental lexicon and that the effects of syllable structure are separable from shared segmental overlap. The data also indicate that syllable representations correspond to a *surface* syllable rather than an abstract underlying syllable posited by some linguistic theories. These findings raise questions concerning the origin of syllable structure in lexical representations. A connectionist simulation utilizing the TIMIT data base shows that syllable-like structure may be induced from exposure to phonetic input. Taken together these results suggest that knowledge of surface syllable structure is actively used in understanding language and this knowledge may derive from a speaker's experience with language.

Introduction

An important enterprise in psycholinguistic research is determination of the lexical properties which underlie our knowledge of language. Early work in this area has identified the importance of semantic relatedness in the organization of the mental lexicon (Neely, 1977). Less well understood is whether phonological properties factor in this organization. Studies indicate that when a subject is asked to determine whether the second member of a word pair is a well formed English word, reaction times are significantly shorter if the preceding word shares some phonological similarity (e.g. MAKE/BAKE vs. RUN/BAKE) (Meyer, Schaneveldt, & Ruddy, 1974; Hillinger, 1980; Jakimik, Cole & Rudnicky, 1985; Slowiaczek, Nusbaum, & Pisoni, 1987; and Emmorey, 1987.) However, the exact locus of these priming effects remain unclear. Two competing factors are implicated in phonological form based priming: First, the amount of phonological overlap (e.g. word pairs like BLAND/BLACK where three segments overlap are more likely to prime than BLEED/BLACK where only two segments overlap). Second, the structure of the overlap (e.g. words pairs like BA.LOON/SA.LOON, where final syllables overlap, are more likely to prime than BREA.KING/SMI.LING, where the final syllables are phonetically different). Unfortunately in past studies these two factors, amount of phonological overlap and the structure of the overlap, have been confounded. The present study was designed to disentangle the effects of shared segmental and syllabic overlap by directly pitting segments and syllables against one another while holding constant the absolute number of shared phonemes.

Experiment 1. Syllable vs. Segment Priming

To explore the effects of shared syllabic and segmental overlap, we compare priming effects in two groups of words; word pairs which share syllabic overlap (either an initial syllable; PAM.PER/PAM.PHLET, or a final syllable; DU.RESS/CA.RESS) and words which share only segmental overlap (e.g. STACK/STAB and BLIS.TER/BLIZ.ZARD). In all cases phonological overlap is approximately three phonemes.

Comparing magnitude of priming for these groups of words permits us to systematically factor out priming effects arising from the quantity of segmental overlap from priming arising from the structure (i.e., syllable structure) of this overlap. Specifically, if it is simply the amount of shared segmental overlap which determines phonological form-based priming, then monosyllabic words such as STACK/STAB should show greater priming than the bisyllabic words which share a syllable (PAM.PER/PAM.PHLET & DU.RESS/CA.RESS). Note that in monosyllabic words, approximately 3/4 of the segments are identical whereas in the bisyllabic words only approximately 1/2 of their total segments overlap. If on the other hand it is the structure of the overlap which is important in phonological priming, we expect greater priming for bisyllabic words which share an initial or final syllable relative to monosyllabic words which do not share syllabic overlap. Importantly, in each case, the amount of segmental overlap is approximately three phonemes. Finally in this last comparison there is a possible source of confound. Specifically, if greater priming is found for the bisyllabic words relative to the monosyllabic words, we cannot be sure the whether observed priming is uniquely attributable to the shared syllables or rather some independent property of bisyllabic words, such as sheer acoustic duration. That is, since bisyllabic words are acoustically longer they might show more robust priming than the shorter monosyllabic words. To control for this confound, we include bisyllabic words which share only a consonant cluster and a vowel but not an entire syllable (e.g. BLIS.TER/BLIZ.ZARD). The inclusion of these word forms will permit examination of the acoustic duration factor in the present experiment. To summarize, this experiment is designed to determine whether the amount of segmental overlap or the structure of the overlap (i.e. syllable structure) is critical in phonological form-based priming. We evaluate phonological priming using an auditory lexical decision paradigm.

Method. Eighteen college students listened to auditorily presented word pairs with an I.S.I. of 100 msec. in an acoustically controlled room. Subjects were to decide the lexical status of the second word of each pair. Subjects pressed one of two computer keys to indicate their choice. Reaction time to respond measured from the end of the second word constituted the dependent variable. For each condition stimuli lists consisted of 18 related-pair trials, 18 unrelated-pair trials and 18 filler items. To determine if priming effects are present reaction times to related-pair trials (e g. STACK/STAB) are compared to unrelated-pair trials (e.g. TRIM/STAB).

Shared Syllables		Examples
Initial	35.7*	PAM.PER/PAM.PHLET
Final	60.0*	DU.RESS/CA.RESS
Shared Phonemes		Examples
Monosyllabic	59.1*	STACK/STAB
Bisyllabic	- 6.1	BLIZ.ZARD/BLIS.TER

Priming in msec. (* $p < .01$)
Table 1

Results. The results presented in Table 1 indicate significant priming for words which shared syllable overlap (both initial and final). In addition, segment priming was found only for monosyllabic words, bisyllabic words which shared an initial consonant cluster and a vowel (but not an entire syllable) did not show priming. The results are consistent with a spreading activation model of lexical access in which both segments and syllables are overtly represented. Moreover the data suggest that higher order syllable representations permit propagation of activation over time. Note for example, we find significant priming for bisyllabic words which share an initial syllable (PAM.PER/PAM.PHLET) but not for bisyllabic words which overlap only in segments (BLIS.TER/BLIZ.ZARD). This result indicates that the syllable has a status independent of the absolute number of shared phonemes. Two important findings emerge from this study: 1) The results support a model of lexical representation in which syllable structure is overtly represented in the lexicon. 2) The results reveal that the effects of shared syllable overlap are separable from shared segmental overlap. However questions remain as to the exact nature of these syllable representations.

Experiment 2. Underlying vs. Surface Syllable Priming

We may make a distinction between surface syllabification and underlying syllabification. Consider for example the word "pony". In slow careful speech speakers syllabify this word as /po.ni/, reflecting perhaps the underlying syllable boundaries. However in fast, everyday speech, we might represent the syllabification as /pon.ni/ where the /n/ appears to be a member of both first and last syllable. This parse is a reflection of surface level syllable structure. Recent work in linguistic theory has greatly elaborated the differences between underlying syllable structure and surface syllable structure. Two differences which factor the construction of the stimuli used in the present experiment include: 1) scope of syllabification and 2) the degree of internal constituency. These differences provide a basis for determining whether the syllable priming observed in Experiment 1 is a reflection of surface or underlying syllable representations.

The scope of syllabification differs for underlying and surface syllables. It has been argued that the domain of syllabification for underlying syllables is limited, whereas surface syllabification is considered exhaustive. The "smaller" underlying syllable serves as a constraint which interacts with word formation processes (Borowsky, 1989). A second difference concerns the presence or absence of internal syllable constituency. Internal constituency refers to hypothetical sub-units which comprise the syllable, these include the onset, nucleus, rime, and coda. Their is considerable evidence from speech production and speech planning for the importance of these constituents in surface syllables (Fromkin 1971, Stemberger 1985; Yaniv, Meyer, Gordon, Huff, & Sevald, 1990). However there is far less evidence for syllable constituency in *underlying* syllable representations (Clements & Keyser 1983).

With these two differences in mind we can construct stimuli sets which contrast surface and underlying syllable structures. Table 2 illustrates underlying and surface representations of the word pairs FAKE/FATE and FAKE/BAKE. In monosyllabic words with long vowels, final segments are extrametrical at the level of underlying syllable structure (Meyers, 1987; Borowsky, 1989). In the diagrams, long vowels are represented as a double occurrence of /a/ in keeping with current phonological theory. Note that in the underlying representations, only the initial consonants and the vowel fall under the scope of the syllable, final consonants are excluded from this domain on the basis of extrametricality (hence marked "ex." in the diagram).

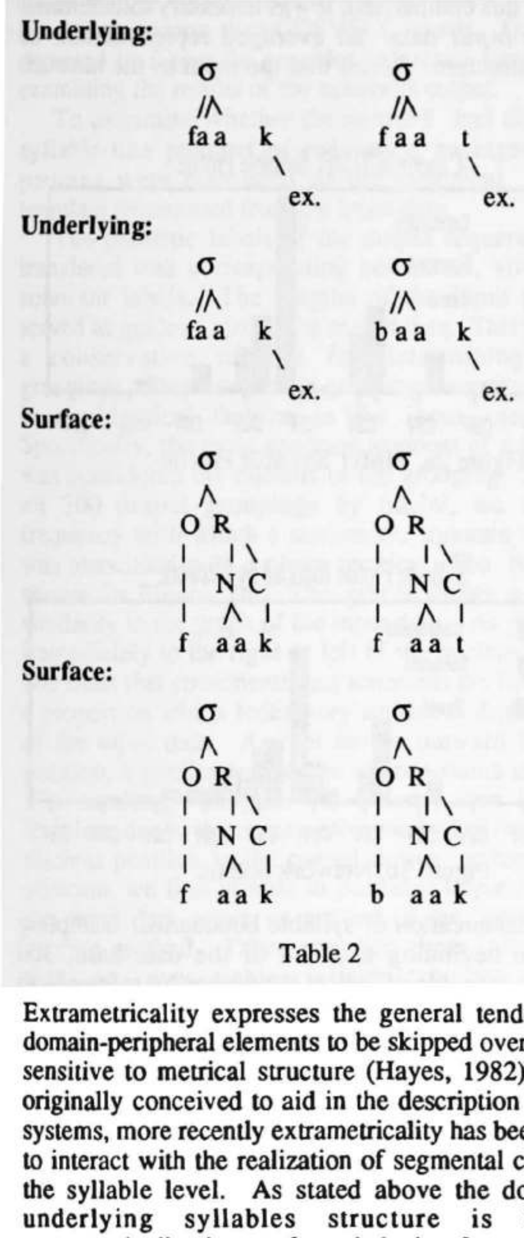

Table 2

Extrametricality expresses the general tendency for domain-peripheral elements to be skipped over by rules sensitive to metrical structure (Hayes, 1982). While originally conceived to aid in the description of stress systems, more recently extrametricality has been shown to interact with the realization of segmental content at the syllable level. As stated above the domain of underlying syllables structure is limited, extrametricality is one formal device for expressing limitations on underlying syllable structure. Given these underlying representations, the words FATE and FAKE will share the underlying syllable /faa/, whereas the words FAKE and BAKE differ in their initial underlying syllables (e.g. /faa/ and /baa/ respectively).

In contrast, in the surface representations all phonemes are exhaustively syllabified and internal constituency is represented. (e.g. onset (O), rime (R) nucleus (N) and coda (C). While neither word pair shares and entire syllable, the word pair FAKE/BAKE, do share the internal rime constituent /aak/, contrariwise the pairs FATE/FAKE lack this relationship.

We predict that if underlying syllabic structure is being primed then the word pairs sharing an entire underlying syllable (e.g., FAKE/FATE) should show a priming effect greater than that of the words which only share partial segmental overlap (e.g., FAKE/BAKE). On the other hand, if surface level syllabification is contributing to priming, we predict that the words FAKE and BAKE may show greater priming due to their shared surface syllable constituency. In summary, comparing priming in words sets which share these characteristics (shared underlying syllables versus shared surface syllable constituents) we may assess whether syllable priming observed in Experiment 1 owes to the activation of surface level or underlying syllable representations.

Method. The data set consisted of ten monosyllabic target words. Three different primes were constructed for each target word. In one case the primes shared underlying syllabic structure (e.g., FATE/ FAKE) in another case, the primes were rhyming pairs sharing surface syllable constituents (e.g., BAKE/FAKE). These two classes of relatedness were compared to phonologically unrelated primes (e.g., WIPE/FAKE). Method and subjects were the same as in Experiment 1.

Underlying Syllable	28 (n.s.)	FATE/FAKE
Surface Syllable	54*	BAKE/FAKE

Priming in msec. (*p < .02)

Table 3

Results. The results shown in table 3 reveal that words which share a surface syllable constituent (i.e. rime) showed significant priming effects, while words which shared underlying syllabic structure did not. These findings suggest that the locus of syllabic priming observed in lexical decision experiments derives from surface rather than an abstract underlying syllabic structure.

Discussion These two experiments argue for a model of lexical representation in which surface syllable structure is overtly represented. These findings raise questions concerning the origin of syllable structure in lexical representations. This issue was explored in a simulation which examined whether syllable like representations could be derived from surface level phonetic input.

Simulation 1.
Induction of Syllable Structure.

The present simulation examines whether syllable-like structure is derivable from naturalistic, phonetically transcribed speech. The simulation uses as input a large data base constructed for studies of automatic speech recognition and uses a neural network to predict structural regularities in the data base. The simulation is highly successful in illustrating the extraction of syllable-like structure from a natural language corpus.

Data The data for the simulation is a subset of the TIMIT data base (Zue, Seneef, & Glass, 1990). The data used in the simulation is derived from the phonetic transcription of the TIMIT sentences provided with the data base. The coding scheme identifies 62 distinct speech sounds and includes demarcations of pauses and ends of sentences. A subset of entire data base was used for the simulation. Ten sentences from 77 randomly chosen male speakers were used, yielding a total of 770 sentences. The phonetic transcription of these sentences was concatenated and arranged sequentially, one phonetic label to a line. All sentence boundary information and pauses were removed from the data set. This yielded a total data set of 27,689 phonetic labels. Importantly, the input data was continuous, no information about word or syllable boundary information is represented in the data set.

Method. A sequentially recurrent network was used in a prediction task as outlined in Elman (1989a). In this case, a sequential network's task is to take successive phonemes from the input and to predict the subsequent phonemes on the output layer. After each phoneme was input, the output was compared with the actual next phoneme, and the back propagation of error learning algorithm (Rumelhart, Hinton, & Williams, 1986) was used to adjust the weights. Localist encodings of the 59 phonetic labels were presented in order with no breaks between words or syllables.

The network (shown in figure 1) consisted of 299 nodes configured to accept 59 inputs and 59 outputs Input was fed to an intermediate layer of 30 units, which in turn was passed to a recurrent layer with 90 units. This was fed to another layer of 30 units and finally back out to the 59 unit output layer. The network was trained through 15 passes through the corpus yielding a total summed squared error of .9278.

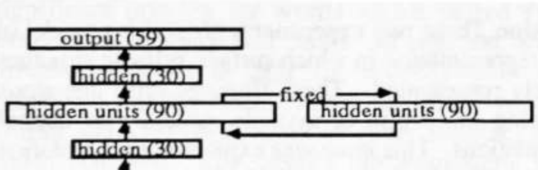

Figure 1. Sequential Recurrent Network

As discussed in Elman (1989), the prediction task is non-deterministic, and short of memorizing the sequences, the network cannot succeed in exact predictions. It is only by virtue of inherent underlying regularities in the corpus that such predictions can be made. Thus, the prediction task provides an avenue by which to discover the regularities of the data set. . The question of interest here is the extent to which this structure may correspond to canonical syllable structure.

Results. To examine the network's success in discovering syllable structure, the extant structure of the input was compared to that predicted by the network. To make this comparison, it was necessary to determine from the *input* data an averaged representation of syllable structure. Recall that the input to the network

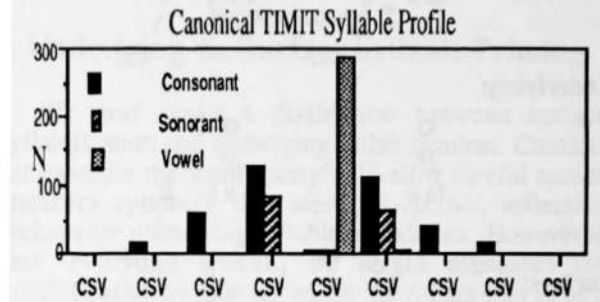

Figure 2a. TIMIT Syllable Profile

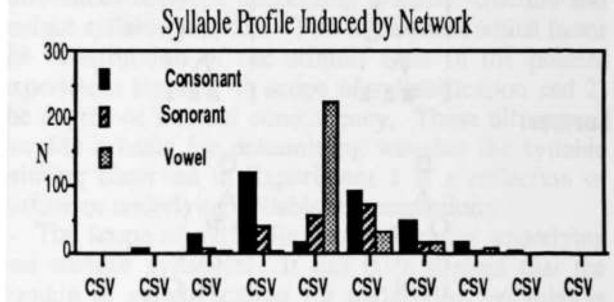

Figure 2b. Network output.

had no demarcation of syllable boundaries. Sampling from the beginning and end of the data base, 300 syllables were identified by the author by reference to the target words. Each phonetic transcription symbol was recoded as belonging to one of three classes of segments: consonants, sonorants, or vowels. The resulting syllables, ranged in length from one to nine segments in length. Next, these 300 syllables were "averaged" to yield the data base's canonical syllable profile. This consisted of determining, for each syllable, the nucleus (typically the vowel) and aligning all 300 syllables according to their nuclei. Finally, counting the number of segment types to the left and right of the nuclei relative to the three classes of segments, provides a "histogram" of canonical syllable preference for the TIMIT data base.

Figure 2a illustrates the canonical TIMIT syllable profile. The graph shows the number of times a particular segment type was found in a position relative to the nucleus. The graph illustrates the preference for English syllables to have a central vocalic nucleus, and the immediate left of the nucleus to be either a consonant or a sonorant. As one moves further from this position, typically only consonants are found. A near reverse profile is found to the right of the nucleus. The diminishing frequency of complex onsets and codas simply corresponds to the fact that CVC words are more common than CCSVSCC words. The profile depicted in figure 2a provides reference template for examining the results of the networks output.

To determine whether the *network* had discovered syllable-like patterns of regularity, averaged output patterns were compared to the canonical syllable template determined from the input data.

The phonetic labels of the output sequences were translated into corresponding consonant, vowel, and sonorant labels. The lengths of the input syllables served as guides to parse the output data. This provided a conservative method for determining output groupings. These sequence groupings were "averaged" in an identical fashion to the input sequences. Specifically, the most sonorant segment of a sequence was considered the nucleus of the grouping. Aligning all 300 output groupings by nuclei, we tally the frequency with which a consonant, sonorant or vowel was associated with a given position. The results are shown in Figure 2b. The graph shows a striking similarity to the graph of the input data. As one moves immediately to the right or left of the nucleus position, one finds that consonants and sonorants are favored, in a proportion which looks very similar to that observed of the input data. As one moves outward from this position, a greater percentage of consonants are found. The resulting "template" is broader than the input template, due to the conservative method of determining nucleus position in the parsed output groupings. In addition, we find vowels in positions beyond those of the input data, again an artifact of the conservative scoring method. Taken together, these facts suggest that an even tighter syllable template could be produced from a less conservative method of analysis. However based on this analysis, it appears that the network, through prediction, is able to extract regularities which bear remarkable similarity to syllable structure regularities observed in the input data. It is important to emphasize that the input data was continuous, with no demarcation of syllable or word boundaries yet the inherent regularities of segment position was sufficient for a network to induce syllable-like structure from exposure to positive instances of data.

Summary. We have examined the ability of a network to extract regularity from a phonetically coded English language data base. The output data revealed that the network was able to extract predictable structure which corresponded, in a striking manner, to syllable structure. The simulation provides extremely strong evidence for the ability to extract syllable structure from positive instances of phonetically labeled data in a natural language data base.

Discussion

Data from two experiments was presented which revealed that syllable structure is used in lexical access. Importantly the data support a model of lexical representation in which facts about surface syllable structure, rather than an abstract underlying syllable structure, is represented. This finding raises questions concerning the origin of this knowledge. A connectionist simulation was presented which demonstrated that a simple learning mechanism, when exposed to a natural language data base, was successful in uncovering syllable-like information. This simulation in part mirrors the child's experience with language. Specifically, we observe that the structure is implicit in the corpus and also in the input to the child. The child's task, like that of the network, is to discover these inherent regularities. However for the model, the input data, while continuous, is nevertheless segmented into phonetic labels. This segmentation gives the model an obvious head start which is not available to the child. Recent work using speech spectrograms as input to a recurrent connectionist network has demonstrated some success in segmentation of speech. (Doutriaux & Zipser 1990; see also Elman 1989b). We may conjecture that knowledge induced from inherent regularities in language provides one basis of organization for the mental lexicon and this knowledge serves in the recognition of words. Taken together these results suggest that the knowledge of syllable structure that a speaker actively utilizes in understanding language may be derived from that speaker's experience with the language. These results have important implications for models of lexical representation.

Acknowledgments

This work was supported in part by Andrus Gerontology Center Grant NIA 5T32AG 00037. I would like to thank Jeff Elman, Liz Bates and Mark Seidenberg for helpful comments on various versions of this work. Alan Petersen and Kim Daugherty also provided valuable input.

References

Borowsky, T. 1989. Structure preservation and the syllable coda in English. *Natural Language and Linguistic Theory*, 7(2):145-166.

Doutriaux, A & Zipser, D. 1990. Unsupervised Discovery of Speech Segments Using Recurrent Networks. In *Connectionist Models, Proceedings of the 1990 Summer School*; D. Touretzky, J.L. Elman, T.J. Sejnowski, and G.E. Hinton Eds. Morgan Kaufmann, San Mateo, CA.

Elman, J. L. 1989a. Representation and structure in connectionist models. CRL Technical Report, 8903, Center for Research on Language, Univ. of California, San Diego.

Elman, J. L. 1989b. Connectionist Approaches to Acoustic/Phonetic Processing. In William Marslen-Wilson (ed.) *Lexical representation and process* MIT Press/Bradford Books, Cambridge MA.

Emmorey, K. D. 1987. Morphological structure and parsing in the lexicon. Ph.D. diss., Univ. of California, Los Angeles, CA.

Hayes, B. 1982. Extrametricality and English Stress. *Linguistic Inquiry*, 13(2):227-276.

Hillinger, M. L. 1980. Priming effects with phonemically similar words: The encoding-bias hypothesis reconsidered. *Memory and Cognition*, 8(2):115-123.

Jakimik, J., Cole, R. A., & Rudinicky, A. I. 1985. Sound and spelling in spoken word recognition. *Memory and Language*, 24:165-178.

Meyer, D. E., Schvanevldt, R .W., & Ruddy, M. G. 1974. Function of graphemic and phonemic codes in visual word recognition. *Memory and Cognition*, 2:309-321.

Meyers, S. 1987. Vowel Shortening in English. *Natural Language and Linguistic Theory*, 5(4):485-518.

Neely, J. H. 1977. Semantic priming and retrieval from lexical memory: Roles of inhibition less spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, 106:226-254.

Rumelhart, D. E., Hinton, G., & Williams, R. J. 1986. Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of Cognition. Volume 1: Foundations.* Cambridge, MA: MIT Press/Bradford Books.

Slowiaczek, L., Nusbaum, H. C., & Pisoni, D. 1987. Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning Memory and Cognition*, 13(1):64-75

Slowiaczek, L., & Pisoni, D. 1986. Effects of phonological similarity in auditory word recognition. *Memory and Cognition*, 14:230-237.

Yaniv, I., Meyer, D., Gordon, P., Huff, C. A., & Sevald, C. A. 1990. Vowel similarity, connectionist models, and syllable structure in motor programming of speech. *Journal of Memory and Language*, 29:1-26.

Zue, V., Seneef, S., & Glass, J. 1990. Speech database at MIT: TIMIT and Beyond. *Speech Communication*, 9:351-356.