

Predicting Noun and Verb Latencies: Influential Variables and Task Effects

Natalie Kacinik (kacinn01@student.ucr.edu)
Christine Chiarello (christine.chiarello@ucr.edu)

University of California, Riverside
Life Sciences Psychology Building
Riverside, CA 92521-0426, U.S.A

Abstract

Natural language comprehension involves processing a multitude of words that vary along many dimensions, some of which may reflect statistical regularities in language. These variables may differ in their relative importance across various types of words and tasks. This study used a multiple regression approach to investigate potentially important predictors of noun-verb processing across naming, grammatical categorization, and sentence completion tasks. Although there were some indications of different predictors for nouns vs verb processing, the strongest predictors of response latencies were primarily determined by the types of processing most important for a given task. One variable of particular interest was the newly created Noun-Verb Distributional Difference (NVDD) metric developed by Chiarello et al. (1999). NVDD values reflect statistical regularities in language regarding the typicality of the contexts in which nouns and verbs tend to occur. The results suggest that although noun-verb typicality, as assessed via the NVDD, is a valid measure of regularities in noun-verb contexts within a linguistic corpus, individuals may not be very sensitive to this dimension in standard psycholinguistic processing tasks.

Introduction

Single word recognition is a central component of language processing. The typical approach has been to use a naming or lexical decision (LD) task and a factorial design to investigate the processing effect of one or more variables such as familiarity or imageability while holding other potentially confounding variables constant. In addition, most single word recognition research has tended to use words of different parts of speech without considering grammatical class (e.g., nouns vs. verbs), or has focused on concrete, imageable nouns. Natural language comprehension, however, involves processing a multitude of words varying along many dimensions. These dimensions may be relatively more or less important for various word types, and their relative importance is likely to vary across different forms of language processing (e.g., word pronunciation vs grammatical identification vs sentence integration).

With a few exceptions (e.g., Balota & Chumbley, 1984; Balota, Cortese, & Pilotti, 1999), there have been few attempts to investigate the relative importance of various orthographic and semantic dimensions for responding to

words across tasks using multiple regression procedures. This approach provides the opportunity to study many variables simultaneously, to determine which lexical dimensions account for the greatest amount of variance in reaction time (RT) and accuracy for a particular task, and to assess whether the variance accounted for is unique, or is shared by other variables.

Such a regression approach was used in the present study to investigate the relative importance of different lexical dimensions across three language tasks. To our knowledge no prior regression study has examined whether various predictor variables are equally applicable to words of different grammatical class. This is an important issue because neuropsychological research has shown that nouns and verbs appear to be processed differently in the brain (e.g., Daniele et al., 1994; Koenig & Lehmann, 1996; Sereno, 1999). It is unclear whether these differences are due to neurally separate noun and verb processing systems, or whether these processing differences are mainly due to different semantic dimensions that covary with word class. Investigating several potentially relevant dimensions using a regression approach may be informative regarding these processing differences between nouns and verbs.

One possible reason word recognition research has generally been limited to concrete, imageable nouns, is the lack of word norming corpora available for other word types. A recent study by Chiarello, Shears, and Lund (1999), however, provides imageability ratings, frequency values from the Unset text corpus of the Lund and Burgess (1996) Hyperspace Analog to Language (HAL) model, and a new measure of noun-verb distributional typicality (the Noun-Verb Distributional Difference, NVDD, metric), for a set of 1197 words: 555 “pure” nouns, 427 “pure” verbs, and 215 words “balanced” for noun-verb usage, as classified by the Francis and Kucera (FK, 1982) norms.

Noun-Verb Distributional Typicality

The new measure of noun-verb usage developed by Chiarello et al. (1999) uses context vectors from the Lund and Burgess (1996) HAL model, where words occurring in similar phrasal and sentential contexts are nearby in high dimensional context space. Context distances were computed between each word and each of the 555 “pure” nouns (according to Francis & Kucera, 1982) and averaged to get a mean noun context distance score. Mean verb distance scores were similarly obtained by computing and

averaging the context distances between each word and the 427 “pure” verbs (Francis & Kucera, 1982). The mean verb context distance was then subtracted from the mean noun context distance, for each word, resulting in a measure referred to as NVDD (Noun-Verb Distributional Difference), indicating the extent to which the word occurs in contexts that are more typical of nouns or verbs. Chiarello et al. (1999) validated the NVDD measure by demonstrating their strong correlation with Francis & Kucera (1982) estimates of noun-verb usage, as well as by examining actual part of speech occurrences for a subset of words in sentence contexts from the Usenet corpus.

The NVDD measure is purely computational, however, reflecting statistical regularities of noun-verb usage in the language corpus. Given the many compelling demonstrations of how information about statistical regularities in a learner’s environment can be extracted to result in the learning of various language phenomena (e.g., Elman, 2001; Saffran, 2001a,b), it was of interest to determine whether the NVDD is psychologically relevant for language processing. We therefore examined whether individuals would demonstrate a sensitivity to the contextual typicality of nouns and verbs in their performance on psycholinguistic tasks. The present paper further investigates the psychological validity of NVDD across three different linguistic tasks, whose findings can be compared with our previous lexical decision results.

Prior Lexical Decision Results

Kacirik, Shears, and Chiarello (2000) reported regression results investigating the influence of imageability (ease with which a word arouses a mental image), word length, experiential familiarity, NVDD, and 2 measures of frequency (FK, 1982, and Usenet corpus) on noun and verb lexical decision (LD) response times (RTs).

The results indicated that noun-verb typicality (NVDD) accounted for a significant portion of verb RT on its own, but it did not account for any unique variance with the other variables in the model. For nouns, the opposite occurred, such that NVDD did account for a small, but significant portion of unique variance. These LD results only partially support the psychological relevance of the recently developed NVDD metric, because it was not a very important predictor. LD, however, requires discriminating words from nonwords, and as nouns and verbs are both words, noun-verb typicality may not really be a relevant dimension for making this discrimination. Experiments 2 and 3 of the current study investigated grammatical categorization and sentence completion tasks, both of which should involve language processes for which noun-verb typicality could be more relevant.

The LD results also showed different variables to be more or less important for predicting noun vs. verb RT. Specifically, imageability appeared to be a more important predictor for verb responses, whereas frequency appeared to be more important for nouns. Most importantly, however, the results indicated that although frequency, imageability,

and NVDD could each account for a portion of the RT variance individually, they failed to account for much unique variance.

The biggest predictor of RT was familiarity, individually accounting for 50% or 60% of the variance, for nouns and verbs, respectively. Moreover, when contributions of the other variables were partialled out, around half of the RT variance accounted for by familiarity appeared to be unique (33% for nouns and 30% for verbs). Familiarity’s importance in predicting LD RTs is not surprising (e.g., Balota et al., 1999; Gernsbacher, 1984), suggesting that familiarity is probably the most important dimension for discriminating between words and nonwords. More surprising, however, was that about half of the variance accounted for by familiarity was unique.

To continue examining predictors of noun-verb processing, and the psychological relevance of the NVDD metric across various linguistic processes, Experiment 1 examined word naming, Experiment 2 involved noun-verb decision, and Experiment 3 investigated sentence completion.

Experiment 1

This experiment investigated which variables would be the most important predictors for noun and verb naming latencies. Word naming entails activating phonological representations to produce a vocal pronunciation response, and is not thought to require much semantic processing (Balota et al., 1999). This is in contrast to lexical decision, as well as the grammatical categorization and sentence completion tasks examined in Experiments 2 and 3, which mainly involve activating semantics to make a decision and subsequent key press response. Familiarity, frequency, and length were expected to be most influential because of their likely importance in the initial recognition processes involved in activating phonological representations (Balota et al., 1999). In addition, because nouns and verbs primarily differ in meaning and grammatical class, we did not expect differences for the relative importance of predictor variables in this pronunciation task.

Method

Mean latencies from Spieler and Balota’s (1997) young adult naming corpus were obtained for 251 nouns and 131 verbs found in our database of 1197 words. Spieler and Balota report that these mean RTs were obtained from 31 Washington University undergraduates (mean age = 22.6), who named a total of 2870 monosyllabic words.

The word length, NVDD, FK and Usenet frequencies, imageability, and familiarity values for each of these words were taken from the Chiarello et al. (1999) database. These 6 predictor variables were combined with the Spieler and Balota naming latencies as the dependent variable.

Results and Discussion

The multiple regression results for noun and verb naming latencies are shown in Table 1. When all 6 variables were in

Table 1: Noun and verb regression analyses for the naming task, Experiment 1.

NOUNS	<u>Sole Predictor</u> <u>Variance</u>	<u>Unique Variance</u> <u>(semi-partial r^2)</u>	<u>Beta</u>	<u>t value</u>	<u>t-test sig. of β</u> <u>$p <$</u>
NVDD	.00	.00	-.00	-0.08	ns
Image	.01	.00	-.01	-0.11	ns
Length	.11	.12	.35	6.08	.0001
FK Freq	.06	.01	-.24	-1.69	ns
Usenet Freq	.05	.00	.09	0.64	ns
Fam	.05	.03	-.19	-2.78	.01
VERBS	<u>Sole Predictor</u> <u>Variance</u>	<u>Unique Variance</u> <u>(semi-partial r^2)</u>	<u>Beta</u>	<u>t value</u>	<u>t-test sig. of β</u> <u>$p <$</u>
NVDD	.01	.02	.17	1.67	ns
Image	.00	.00	-.01	-0.15	ns
Length	.07	.05	.23	2.52	.05
FK Freq	.02	.00	.04	0.33	ns
Usenet Freq	.03	.00	-.10	-0.73	ns
Fam	.05	.03	-.25	-2.02	.05

Note: $R^2 = .21$, $F(6, 245) = 10.57$, $p < .0001$, for nouns, and $R^2 = .12$, $F(6, 121) = 2.87$, $p < .05$, for verbs

the regression, they significantly accounted for 21% of the RT variance for nouns, and only 12% of variance for verbs. It thus appears that the lexical dimensions typically thought to influence word naming are relatively more important for nouns than verbs. Further examination of the results, however, suggests that although both frequency measures appear to be additionally influential for nouns, they do not account for any unique variance. The variance they account for appears to be subsumed by length and familiarity. Indeed, the main conclusion from these results is that similar dimensions (length and familiarity) were most important predictors for both noun and verb RTs. These findings are in contrast to the previous LD results where, after familiarity, there was some indication that different predictors were important for noun vs verb processing. This suggests that noun-verb processing differences are due to semantics and/or result from postlexical processing.

NVDD was not found to be a significant predictor for noun latencies, and was only marginally significant for verbs, accounting for 2% unique variance. Thus, both the LD and naming results provide minimal support that individuals are sensitive to regularities in the contexts in which nouns and verbs occur. Another possibility is that neither of these tasks involves explicitly activating word class information. Perhaps a task that does require explicit activation of noun or verb meaning and/or grammatical class information, will show greater effects of noun-verb typicality.

Experiment 2

This experiment investigated the relevance of the typicality of contexts in which nouns and verbs tend to occur (NVDD) for deciding whether a word is a noun or verb. In contrast to LD and naming, this task requires the explicit activation of

grammatical class information in order to make the noun-verb decision.

Method

Participants

Forty native English speaking University of California, Riverside undergraduates (20 males) participated in the experiment in exchange for course credit or pay (\$6.00).

Stimuli

The same set of 152 nouns and 137 verbs, varying in NVDD, from Kacirik et al. (2000) were employed here.

Procedure

Each trial began with the presentation of a 400 ms fixation point, followed by 100 ms blank screen, which was followed by presentation of the target word. Participants were required to decide whether each item was a noun or a verb by making a button press response as quickly as possible. Targets remained on the screen until they responded, and the inter-trial interval was 1500 ms. Participants were told that nouns were words naming a quality, person, place, or thing, while verbs are words that express an action or the occurrence of an event, and given some examples of each. Fifteen practice trials preceded the experiment.

Results and Discussion

The noun and verb multiple regression analyses are presented in Table 2. When all the variables were included in the regression model, it was better at predicting noun-verb decision than naming latency, such that 44% of the RT

Table 2: Noun and verb regression analyses for the noun-verb decision task, Experiment 2.

NOUNS	<u>Sole Predictor</u> <u>Variance</u>	<u>Unique Variance</u> <u>(semi-partial r^2)</u>	<u>Beta</u>	<u>t value</u>	<u>t-test sig. of β</u> <u>$p \leq$</u>
NVDD	.04	.03	.19	2.87	.005
Image	.37	.28	-.57	-8.50	.0001
Length	.02	.01	.12	1.82	ns
FK Freq	.01	.00	-.01	-0.06	ns
Usenet Freq	.01	.00	-.10	-0.89	ns
Fam	.05	.01	-.11	-1.43	ns
VERBS	<u>Sole Predictor</u> <u>Variance</u>	<u>Unique Variance</u> <u>(semi-partial r^2)</u>	<u>Beta</u>	<u>t value</u>	<u>t-test sig. of β</u> <u>$p \leq$</u>
NVDD	.10	.07	-.28	-3.79	.001
Image	.24	.17	-.44	-6.08	.0001
Length	.00	.00	.01	0.09	ns
FK Freq	.01	.01	.11	1.16	ns
Usenet Freq	.02	.01	-.12	-1.30	ns
Fam	.18	.03	-.20	-2.40	.05

Note: $R^2 = .44$, $F(6, 144) = 18.57$, $p < .0001$, for nouns, and $R^2 = .40$, $F(6, 123) = 14.39$, $p < .0001$, for verbs

variance for nouns, and 40% of the RT variance for verbs, was accounted for.

For nouns, imageability followed by NVDD was the most significant predictor of noun-verb decision times. Although familiarity individually accounted for 5% of the noun RT variance, it was not found to contribute a significant amount of unique variance. Imageability and NVDD were also the most important predictors of unique variance for verbs. Familiarity was also a strong predictor of verb latency on its own, but in contrast to the noun results, it also accounted for a small significant unique amount of variance.

Imageability therefore appears to be the most influential dimension for deciding whether a word is a noun or a verb. Furthermore, much of imageability's contribution and the variance it accounts for seems to be unique. There is also some indication that imageability is more important for nouns than verbs in the noun-verb decision task. This is contrary to the LD results, where imageability appeared somewhat more important for verbs. A possible explanation for these results is that since nouns are generally more imageable than verbs, a highly imageable concrete word encountered in the context of the noun-verb decision task must be a noun and could be responded to very rapidly (Chiarello et al., 1999). A negative correlation should thus be expected between imageability and decision latencies for verbs, because low imageability words are more likely to be verbs (Chiarello et al., 1999), and should be responded to quickly. The obtained correlation, however, was positive, suggesting that imageability does not facilitate the noun-verb decision per se. Instead, we suggest it facilitates earlier processes such as the speed of word meaning activation, enabling subsequent noun-verb decisions to be made more rapidly.

As expected, noun-verb typicality (NVDD) was significantly correlated with both noun and verb decision

latencies ($r = .21$, $p < .01$, and $r = -.32$, $p < .0001$, respectively). Given that noun-verb decision does explicitly involve processing part of speech information, it is surprising that NVDD was not more important and only accounted for 3% and 7% unique variance for nouns and verbs, respectively. Recall, however, that the NVDD is a measure of the typicality of contexts in which nouns and verbs occur. All tasks investigated thus far involved single word processing, and may not reflect the influence of the fundamentally contextual nature of the NVDD metric. It is possible that NVDD may be most relevant for processing words in sentence contexts.

Experiment 3

This experiment investigated whether noun-verb contextual typicality would influence response speed for deciding whether a word could be sensibly integrated into an incomplete sentence. If sentence context is assumed to constrain possible completions, faster latencies would be expected for words that are highly typical nouns or verbs than for words less typical of their grammatical class.

Method

Participants

Forty native English speaking University of California, Riverside undergraduates (20 males) participated in the experiment in exchange for course credit or pay (\$6.00).

Stimuli

Incomplete sentence frames, which could be sensibly completed by either a high or low NVDD noun or verb, were created. For example, *punish* and *smack* are high and low NVDD verb completions for "the father wanted to

_____ his son”, respectively, and *tavern* and *pub* are high and low typicality noun completions for “*They walked into the _____*”. We created such incomplete sentences for 80 verbs (40 high and 40 low) and 88 nouns (44 high and low) from the set of 152 nouns and 137 verbs used by Kacinik et al. (2000). Sentences were normed and balanced for sensibility and cloze probability. Nonsensical completions were created by re-pairing sentence frames and target words (e.g., *She had to punish the text*).

Procedure

Each trial began with the presentation of a 500 ms fixation point, immediately followed by the appearance of the incomplete sentence. After 1200 ms, the target word appeared above the sentence, allowing participants enough time to read the sentence prior to the target’s appearance. Both the sentence and target word remained on the screen until they responded. Participants had to decide whether the target word was a sensible completion to the sentence and responded by pressing a key as quickly and accurately as possible. The inter-trial interval was 1000 ms. Twenty-four practice trials were completed prior to the experiment.

Results and Discussion

Table 3 presents results from regression analyses for the related condition (i.e., when the noun or verb was a good completion to the sentence). Contrary to our predictions, noun-verb contextual typicality (NVDD) was not found to be relevant for sentence integration, a task where it was expected to strongly influence processing. Indeed, it is surprising that with the exception of Usenet frequency for nouns ($r = -.22$, $p < .01$) and imageability for verbs ($r = -.25$, $p < .01$), none of the variables were significantly correlated with decision latencies. Noun RT did not correlate with NVDD ($r = -.18$, ns), imageability ($r = -.20$, ns), length ($r =$

$.07$, ns), FK freq ($r = -.12$, ns), and familiarity ($r = -.11$, ns). Verb RT was also not correlated with NVDD ($r = -.16$, ns), length ($r = .09$, ns), FK freq ($r = -.01$, ns), Usenet freq ($r = -.10$, ns), and familiarity ($r = -.11$, ns). These results suggest that lexical- semantic dimensions identified as being important for single word recognition are much less relevant for integrating words into sentences. Indeed, when all the variables are in the regression model, they only account for a marginally significant (13%) portion of variance for nouns, and a non-significant 12% of the variance for verbs.

One possible explanation is that these are lexical variables, representing characteristics of single words, and this task is primarily measuring sentence integration. The nature of the sentence, therefore, is also an important source of variance for this task. Indeed, it may even be the most important, suggesting that perhaps “sentence-level variables” such as sentence length, imageability, or meaningfulness, would be better predictors of decision latency. Another possibility is that these lexical variables are still important for initial word recognition and meaning activation, but that their influence dissipates once a word has been recognized, such that they are not involved in higher-level sentence integration processes. This would predict these lexical dimensions, and possibly NVDD, to significantly contribute to initial word recognition processes that happen in on-line sentence comprehension.

General Discussion

The relative influence of variables on processing nouns and verbs in naming, noun-verb decision, and sentence completion tasks was investigated. There were some indications in our previous lexical decision results that, after familiarity, different dimensions might vary in terms of their relative importance for noun versus verb processing.

Table 3: Noun-verb regression analyses for the sentence completion task, Experiment 3.

NOUNS	<u>Sole Predictor</u> <u>Variance</u>	<u>Unique Variance</u> <u>(semi-partial r^2)</u>	<u>Beta</u>	<u>t value</u>	<u>t-test sig. of β</u> <u>$p \leq$</u>
NVDD	.03	.03	-.18	-1.57	ns
Image	.04	.05	-.25	-2.24	.05
Length	.00	.00	.01	-0.08	ns
FK Freq	.01	.00	.03	0.18	ns
Usenet Freq	.05	.04	-.28	-1.85	.10
Fam	.01	.01	.11	0.81	ns
VERBS	<u>Sole Predictor</u> <u>Variance</u>	<u>Unique Variance</u> <u>(semi-partial r^2)</u>	<u>Beta</u>	<u>t value</u>	<u>t-test sig. of β</u> <u>$p \leq$</u>
NVDD	.03	.03	-.17	-1.45	ns
Image	.06	.05	-.25	-2.12	.05
Length	.01	.00	.01	0.10	ns
FK Freq	.00	.01	.19	1.04	ns
Usenet Freq	.01	.01	-.20	-1.02	ns
Fam	.04	.00	-.05	-0.38	ns

Note: $R^2 = .13$, $F(6, 81) = 1.96$, $p < .10$, for nouns, and $R^2 = .12$, $F(6, 73) = 1.66$, ns, for verbs

Although the present study also found some evidence that different variables were relatively more important for the processing of nouns vs verbs across tasks, these differences were fairly subtle. They generally involved differences in the strengths of relationships between variables and RT, or in the unique amount of variance accounted for by each variable. Thus, the main conclusion from this study should be that the strongest predictors of RT do not depend on word class per se, but are determined primarily by the type of processing necessary for a given task. Specifically, word length and familiarity were most important for naming both nouns and verbs, familiarity is the most important dimension for discriminating words and nonwords (regardless of whether the item is a noun or verb), imageability and noun-verb typicality were most relevant for deciding whether a word is a noun or verb, and imageability was the most relevant lexical variable for integrating words into sentences.

A variable of particular interest in the current study was the typicality of contexts in which nouns and verbs tend to occur, as measured by Chiarello et al.'s (1999) recent NVDD metric. Despite the fact NVDD captures statistical regularities of noun-verb usage based on the typicality of contexts in which they tend to occur (Chiarello et al., 1999), it does not seem very relevant for the processing of nouns and verbs across a variety of tasks. Some results did find noun-verb typicality to correlate with lexical and noun-verb decision latencies. In these cases, however, part of that correlation with RT was also accounted for by other variables, such that the unique portions of variance accounted for by NVDD were rather small. This was true even in tasks where noun-verb typicality was expected to influence processing.

There is no doubt that individuals are sensitive to some statistical regularities present in the language environment, many of which can affect processing (e.g., Saffran, 2001a). The present findings suggest, however, that this may not be true for all such regularities. The typicality of contexts in which nouns and verbs tend to occur, as measured via NVDD, appears to be a valid regularity within a linguistic corpus. Yet individuals may not be very sensitive to this dimension in standard psycholinguistic tasks.

References

- Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. Journal of Experimental Psychology: Human Perception & Performance, 10, 340-357.
- Balota, D.A., Cortese, M.J., & Pilotti, M. (1999). Item-level analyses of lexical decision performance: Results from a mega-study. In Abstracts of the 40th Annual Meeting of the Psychonomics Society. Los Angeles, CA: Psychonomic Society.
- Chiarello, C., Shears, C., & Lund, K. (1999). Imageability and distributional typicality measures of nouns and verbs in contemporary English. Behavior Research Methods, Instruments, & Computers, 31, 603-637.
- Daniele, A., Guistolisi, L., Silveri, M.C., Colosimo, C., & Gainotti, G. (1994). Evidence for a possible neuroanatomical basis for lexical processing of nouns and verbs. Neuropsychologia, 32, 1325-1342.
- Elman, J. L. (2001). Connectionism and language acquisition. In M. Tomasello & E. Bates (Eds.), Language development: The essential readings (pp. 295-306). Malden, MA: Blackwell.
- Francis, W., & Kucera, H. (1982). Frequency analysis of English usage: Lexicon and grammar. Boston: Houghton Mifflin.
- Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. Journal of Experimental Psychology: General, 113, 256-281.
- Kacirik, N., Shears, C., & Chiarello, C. (2000). Familiarity for nouns and verbs: Not the same as, and better than, frequency. Proceedings of the 22nd Annual Meeting of the Cognitive Science Society, Philadelphia, PA.
- Koenig, T., & Lehmann, D. (1996). Microstates in language-related brain potential maps show noun-verb differences. Brain & Language, 53, 169-182.
- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrences. Behavior Research Methods, Instruments & Computers, 28, 203-208.
- Saffran, J. R. (2001a). The use of predictive dependencies in language learning. Journal of Memory & Language, 44, 493-515.
- Saffran, J. R. (2001b). Words in a sea of sounds: The output of infant statistical learning. Cognition, 81, 149-169.
- Sereno, J.A. (1999). Hemispheric differences in grammatical class. Brain & Language, 70, 13-28.
- Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. Psychological Science, 8, 411-416.

Acknowledgments

We thank Connie Shears for her help in developing the database used in the current study, and Curt Burgess, Dan Ozer, and Robert Rosenthal, for their useful suggestions. This research was supported by National Science Foundation Grants No. SBR-9729009 and BCS-0079456 to the second author.