LACUS FORUM XXXII

Networks

Dartmouth

TOWARDS A SINGLE-MECHANISM ACCOUNT OF FREQUENCY EFFECTS

VSEVOLOD KAPATSINSKI
Indiana University

USAGE-BASED APPROACHES TO LINGUISTICS[1] are committed to explaining language structure in terms of domain-general abilities and mechanisms influencing language use. A central focus of investigation has been frequency effects. However, relatively little attention has been paid to why frequency has the effects that it does. That is, what are the cognitive mechanisms behind sensitivity to frequency? This is the question addressed in this paper.

1. TOKEN FREQUENCY EFFECTS IN LEXICAL ACCESS AND PRIMING. The greater the token frequency of a lexical item, the faster its activation occurs. For word recognition, see Coltheart *et al.* (1977), Glanzer and Ehrenreich (1979), Norris (1984), Paap *et al.* (1987), Goldinger *et al.* (1989), Monsell (1991), Luce *et al.* (2000), and Plaut and Booth (2000).

The second major frequency effect in word recognition is that the higher the token frequency of either the prime or the target, the smaller the amount of priming observed,[2] cf. Scarborough *et al.* (1977), Forster and Davis (1984), Norris (1984), Stark (1997), Perea and Rosa (2000), and Versace and Nevers (2003) for identity priming; Moder (1992) for prime frequency in morphological priming; Thomsen *et al.* (1996) for prime frequency in semantic priming; Goldinger *et al.* (1989) and Luce *et al.* (2000) for prime frequency in inhibitory phonological priming; Schuberth and Eimas (1977), Neely (1991), and Plaut and Booth (2000) for target frequency in semantic priming; and Perea and Rosa (2000) for target frequency in orthographic priming.

Slow recognition leads to low experienced token frequency since slow processors would not recognize as many tokens of a type present in the environment as fast processors, hence Plaut and Booth's (2000) finding that when lexicon size is controlled children who are slow at word recognition show more priming. Conversely, low token frequency of a word leads to slow recognition of the word.

2. PREVIOUS ACCOUNTS. Within Network Theory, Moder (1992) proposed that high token frequency weakens a word's connections to neighboring words. Priming occurs by the spread of activation from the prime to the target. Since high-frequency words have weak connections, they receive less activation from their neighbors and their neighbors receive less activation from them. This account does not explain why high token frequency also reduces identity priming where the prime and the target are the same word.

Ratcliff and McKoon (1988) proposed that the prime and the target form a compound cue used to access long-term memory. The greater the familiarity of the cue, assessed as a weighted sum of the familiarities of the prime and the target, the faster long-term-memory access can occur. The greater the frequency of the target, the smaller the prime's contribution to overall familiarity of the cue, hence the amount of priming observed is smaller when the target is very familiar, i.e. has high token frequency. This account predicts that high prime frequency should increase priming, since high-frequency primes would contribute much to the overall familiarity of the cue, while in reality high prime frequency reduces the magnitude of priming observed.

Plaut and Booth (2000) propose a distributed connectionist model as a way to account for frequency effects. In this model, the prime and the target are overlapping patterns of activation (which can be understood as ordered sets of 1's and 0's) superimposed on the same set of nodes. The more similar the prime and the target are, the more values they will have in common, hence the transition from the prime to the target will be less costly for the network when the prime and the target are similar: fewer changes to node activation values have to be made. Nodes have sigmoid activation functions, hence more input activation is required to change a node's activation level by a fixed amount in the direction of a value (1 or 0) when the node's resting activation level is already close to that value. Hence, activation of a high-frequency node value during prime presentation will not improve the node's ability to take on that value as much as when the value is mid-frequency. This account does not explain why prime frequency and target frequency influence the amount of priming. According to the account, it is only the frequency of the node values that the prime and the target share that should matter, because the prime and the target do not have an independent existence. The importance of whole-unit frequency in priming is shown by priming asymmetries. For a given pair of stimuli, less priming is observed when the high frequency member of the pair is the prime than when it is the target in semantic (Koriat 1981, Chwilla *et al.* 1998), visual (Rueckl 2003), morphological (Schriefers *et al.* 1992, Feldman 2003), acoustic/phonetic (Goldinger *et al.* 1989) and phonological (Radeau *et al.* 1995) priming.

3. RELEVANT FEATURES OF THE LOCAL ACTIVATION SPREAD THEORY (LAST).

3.1. ARCHITECTURE. LAST proposes that memory is a localist associative network where every unit type—a word, a morph, a phone, a construction, a non-verbal stimulus—corresponds to a TYPE NODE, and every presentation of a type forms a TOKEN NODE (as in the weakly abstractionist approach of Bowers 2000).

In LAST, all types are connected to each other, although these connections vary widely in their strength, while a token is linked to only one type. Evidence for full connectivity between types comes from Ratcliff and McKoon (1981), who found that degree of semantic relatedness between the prime and the target influences the magnitude of the priming effect but not how soon after prime presentation the effect can be observed; thus, closely related words and less strongly related ones appear to be equally close to each other. If this were not the case, activation spreading from the prime would take longer to reach distantly related targets than closely related ones. Therefore, more time would need to pass after prime presentation for the effect of the prime's presentation to be observed with distantly related targets than with closely related targets. Since no differences are found, activation must reach distantly related and closely related targets simultaneously.

The architecture is presented in **Figure 1**.

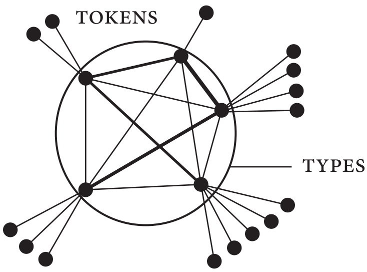

Figure 1. *Architecture of memory according to LAST.*[3]

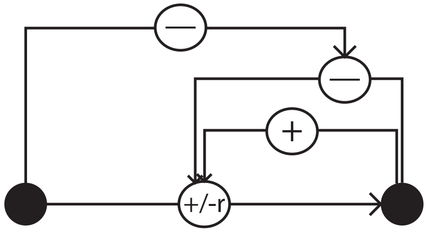

Figure 2. *Link structure in LAST.*[4]

3.2. LINK STRUCTURE. In LAST, a LINK is a unidirectional channel of activation flow in that it only transmits activation from the HEAD of the link to its TAIL. In **Figure 2** a link is represented as an arrow pointing in the direction of activation flow, that is, from the head of a link to its tail. Each CONNECTION in the network consists of two links such that the head of one link is the tail of the other and vice versa. Each link has a PROPAGATION FILTER (PF). Because the two links in a connection have different propagation filters, a connection is represented by two arrows, rather than a single bi-directional one. The resting activation value (r-value) of a link's PF is directly proportional to how much activation is allocated to the link by its head. The PF, however, is not affected by activation spreading through the link it is located on (PF's were first introduced by Sumida and Dyer 1992 and elaborated on in Sumida 1997). If activation flowing through a link increased the PF's r-value, the link would strengthen whenever its head is activated, wrongly predicting that high token frequency words are better linked to their neighbors and therefore are better able to activate or prime them, against findings that high token frequency actually corresponds to reduced priming (see next section).

On the other hand, a link should strengthen whenever its head and tail are activated simultaneously (co-activated) to allow associative learning to occur. This would imply that the PF's r-value is raised whenever the head and the tail of the link are co-activated and hence that the PF of a link is a tail on a subsidiary link (LINKTRON) headed by either the head or the tail of the link the PF mediates. The influence of this linktron must, however, be counteracted whenever the head or the tail is activated in isolation. The link structure shown in **Figure 2** is the simplest one consistent with the evidence thus far (see Kapatsinski 2005 for relevant discussion).

3.3. DYNAMICS OF ACTIVATION SPREAD. LAST, like many other theories of word recognition (e.g. Morton 1969), proposes that each node has an ACTIVATION THRESHOLD. After the activation threshold is reached, activation is divided between the node itself and all links headed by the node. The amount of activation leaving a node is limited and as activation is leaving a node it is divided between all links connected to the node (Anderson 2000). LAST differs from previous semantic-network models in assuming that the node itself participates in this competition so that the more links are connected to a node, the less activation will be allocated to any one link and to the node itself. Finally, LAST proposes the EQUITY PRINCIPLE, which states that the amount of activation allocated to a link is directly proportional to the strength of that link (found also in previous spreading-activation models, cf. Zeelenberg *et al*. 2003 for review). However, while the strength of a link is equivalent to the resting activation level of its propagation filter, the strength of a node, lacking a propagation filter, is fixed. Activation stored in a node or a PF is assumed to decay as time progresses. The decay function is exponential or power-law based (cf. Sikstrom 2002). Decay rate at a point of time is specific to the date of birth and size of a given ACTIVATION UNIT where an activation unit is a moving element defined by its current location as well as its time and place of origin. Activation units created recently decay at a faster rate than those created long ago and the larger the activation unit, the slower its rate of decay.

4. THE LAST ACCOUNT OF THE FINDINGS.

4.1. THE ARCHITECTURE: TOKEN FREQUENCY, LEXICON SIZE, AND SPEED OF PROCESSING. In priming, when the prime is presented, matching types are partially activated. The way this recognition process occurs is by matching the incoming token to already existing tokens (cf. Hintzman 1986, Johnson 1997, Pierrehumbert 2001 for similar proposals) such that more activation is allocated to tokens that are more similar to the incoming token. Since high-frequency types have more tokens, they will be activated more quickly and low-frequency types are more likely to be misrecognized as high-frequency types than vice versa. Whenever a type passes the activation threshold, a token node is created storing the information about that particular instance of the type, and activation starts to spread from the type node. If no type is activated, a new type is formed. Since the spread of activation from the node along type-token links occurs only after it has been activated, speed of the type node's recognition is directly proportional to its r-value: the lower the type node's r-value, the more input activation is needed to reach the level sufficient for recognition and the less input activation is coming in from the tokens. Thus, the greater the token frequency of a type, the faster its recognition should occur.

It has often been observed that logarithmically scaled frequency and not raw frequency has linear correlations with various dependent variables, reaction times in word recognition or lexical decision studies being one example. In LAST, what happens when token frequency of a type is increased by one token, is that the resting activation level of the type is raised and a new type-to-token link is acquired by the type. The more frequent the type, the less activation will remain in the type node, hence the less its resting activation will be increased. In addition, the more frequent the type, the less activation will be allocated to the propagation filter of the newly formed type-to-token link.[5]

Once the prime type is activated, activation starts to flow out from it. As it is leaving the node, it is divided between the node itself and all links headed by the node, one of the links being tailed by the target, which has not yet been presented. Thus, the greater the number of links headed by the prime's type, the less activation will remain in the node and the less activation will be allocated to any one link. Given that every type is connected to all other types while a token is only connected to one type, the only factors influencing the number of links radiating from a type are its token frequency and the number of types in the lexicon. Therefore, the higher the token frequency of the prime, the less priming, including identity priming, should occur. Thus, the same process appears to be implicated in the prime frequency effect in priming as in habituation.

Slow recognition feeds back to low experienced token frequency since slow processors would not recognize as many tokens of a type present in the environment as fast processors and hence would have fewer type>token links per type, resulting in increased priming. In addition, the smaller the size of the lexicon, the more activation is allocated to any one node, hence more priming in children and late signers (as found by, e.g. Simpson & Lorsbach 1983, Emmorey *et al.* 1995, Nation & Snowling 1998, Castles *et al.* 1999), and faster recognition (and higher rated familiarity) of words learned early in life compared to words with the same token frequency learned later in life, indicating higher resting activation levels for early-acquired words (as found by, e.g., Brown & Watson 1987, Bonin *et al.* 2001, and Ghyselinck *et al.* 2004).

The finding that early-acquired words exhibit less identity priming (Barry *et al.* 2001) is predicted because when the lexicon is small, more activation will reach the propagation filter of a link when the nodes it links are co-activated. On the other hand, the proportion of activation received by the type node itself is constant because of the node's lacking a PF. Thus, LAST predicts that early age-of-acquisition should be correlated with having many strong associates, as found by Steyvers and Tenenbaum (2005).

4.2. THE EQUITY PRINCIPLE: NEIGHBORHOOD DENSITY, DENSENSITIZATION, AND BLOCKING. The Equity Principle states that the amount of activation allocated to a link is proportional to its strength (cf. Anderson 2000). If this is the case, we should predict that words that are semantically, phonologically, or orthographically similar to many other words, that is, words located in dense neighborhoods, should exhibit less priming than words located in sparse neighborhoods, since a link or node of a given strength will receive less activation in a dense neighborhood than in a sparse neighborhood. This is indeed what is found by Thomsen *et al.* (1996) for semantic priming and by Perea and Rosa (2000) for orthographic priming.

Even more direct evidence for the Equity Principle is provided by Anaki and Henik (2003), who find that if the target is given as an associate of the prime by a certain percentage of subjects in a free association task, there will be more priming between the prime and the target if the other associates of the prime are given by low percentages of the subjects or if the prime has fewer associates than if other associates are also given by many subjects or many associates exist.

The Equity Principle also explains the blocking effect in associative learning: if an unconditioned stimulus (US) is already strongly associated with a conditioned stimulus (CS1), it is hard to associate with another conditioned stimulus (CS2) (e.g. Kamin 1969, Marchant & Moore 1973). This is because the linktron headed by the US and tailed by the PF of the CS2-US link has to compete with a strong US-CS1 link, relative to the case in which the US has no strong associates. Consequently less activation will be allocated to the CS2-US's PF when US is strongly associated with CS1, leading to slower strengthening of CS2-US.

Similarly, high-frequency CS1's and US1's are harder to associate with a CS2 or US2 because the PF's of links being acquired have to compete with many type-token links headed by CS1 or US1 (what are known as the CS pre-exposure effect and the US desensitization effect, cf. Hall 2003, Baysari & Boakes 2004).

Further evidence for the Equity Principle in associative learning comes from the interference paradigm of Barnes and Underwood (1959) and McGovern (1964), who found that when subjects are asked to learn a list of A–C stimulus pairs after learning a list of A–B pairs, they can recall B when presented with A worse than subjects who are asked to learn C–D pairs after learning A–B pairs.

Finally, the Equity Principle explains transitional probability effects (e.g. the ability to implicitly segment a speech stream into words based on only transitional probabilities present in the speech stream demonstrated by Aslin *et al.* 1998) without postulating that speakers calculate probabilities: a word that is very probable is a word whose connection to the previous word in the processing sequence is strong relative to its competitors.

4.3. SIZE-DEPENDENT DECAY; PERSISTENCE OF PRIMING, PRIMING ASYMMETRIES. A challenge to any single-mechanism account of priming is to explain why identity priming persists over several intervening items, while phonological and orthographic priming decay rapidly (cf. Stockall 2004). LAST handles this dissociation by assuming that the node from which activation spreads receives the lion's share of that activation. Since in LAST, the larger the activation unit, the slower the rate of decay, we predict that identity priming and priming that relies on stronger connections (relative to all connections headed by the source node) should decay less rapidly.

One untested prediction of this account is that identity priming should persist for a longer time if the primed unit is low-frequency than if it is high-frequency, if it has few neighbors than if it has many neighbors, and if it is in a large lexicon rather than a small lexicon. Some evidence for slower decay of activation in small lexicons is provided by the fact that children are commonly found to exhibit more perseveration errors relative to all errors produced than adults on both the phonological and the lexical levels (Stemberger 1989, Jaeger 2004).[6]

Size-dependence of decay rate also allows us to account for priming asymmetries. If larger activation units decay more slowly, less priming should be observed when the prime is high frequency, and the target is low frequency than when the prime is low frequency and the target is high frequency. The reason is that the major reduction in activation unit size occurs earlier in a unit's lifetime when the prime is high frequency than when the target is high frequency. As a result, activation spreading from the prime decays more before target presentation when the prime is high frequency than when the target is high frequency, resulting in decreased priming.

5. CONCLUSION. LAST provides an explicit, testable account of frequency, neighborhood density, lexicon size, age of acquisition, target degradation, and speed of processing effects across tasks, domains, and species, using the single mechanism of spreading activation in a localist associative network, whose structure is independently motivated.[7]

---

1 Many thanks to Joan Bybee, William Sullivan, Rena Torres-Cacoullos, Catherine Travis, and the anonymous reviewers for helpful comments on earlier drafts of this paper and to Larry Gorbet, Dan Jurafsky, James Macfarlane, James McClelland, Jill Morford, Dan Olsher, David Pisoni, and the audiences at LACUS 32 and HDLS VI for stimulating discussions of the issues involved.

2 By 'amount of priming' we mean the number of milliseconds by which the reaction times to a word preceded by a related word differ from reaction times to the same word preceded by an unrelated word. Priming can be excitatory, when reaction times after a related word are faster than after an unrelated word or inhibitory when the reverse is true. The word whose reaction times we are measuring is called a target while the preceding related word is called a prime. The prime and the target may be similar in various ways: they can share acoustic characteristics, phonemes, letters, morphemes, meaning, or everything. Accordingly, we distinguish between phonetic, phonological, orthographic, morphological, and identity priming. On the sentence level, there is also syntactic priming, where the prime sentence and the target sentence share a syntactic construction, e.g., both are passive. Orthographic priming involves visual presentation of the prime and the target while phonological and phonetic priming involve auditory presentation. Morphological priming is often crossmodal, where the prime and the target are presented in different modalities.

3 Lines represent connections, width of line represents connection strength, filled circles represent nodes. The circle surrounding the type nodes has no theoretical significance.

4 If an r-value is negative, excitatory activation entering the node will become inhibitory when it leaves the node, where INHIBITORY activation reduces the r-value of nodes it enters. Linktron PF's are binary and non-trainable. The necessity of these properties of linktron PF's is shown in Kapatsinski (2005). A linktron can transmit activation by default but if any amount of inhibition is applied to its PF it does not transmit any activation. Transmitting activation (or inhibition) to a linktron's PF has no effect on the PF's resting activation level. Thus each linktron headed by a node A receives the same proportion of A's activation.

5 Calculations shown in Kapatsinski (2005) indicate that the r-value of type>token links must depend on the r-value of the type for the appropriate relations between the model's parameters to hold for all r-values of the (source) type.

6 Greater proportion of perseverations in small lexicons can be achieved even if the same decay rate is applied because the decaying activation unit is larger. However, this is unlikely to explain the persistence of identity priming: at a short delay, identity priming is no more than 3 or 4 times larger than similarity-based priming for near synonyms. Yet, identity priming can persist for a

very long time while semantic priming usually decays after one or two intervening items. The reason semantic priming can persist if the task involves much semantic processing (Joordens & Becker 1997) may be because the task draws attention to semantic representations, either allowing them to attract larger activation units, slowing the decay rate, or both.

7 See Kapatsinski (2005) for a LAST account of type frequency and relative frequency effects.

REFERENCES

Anaki, D. & A. Henik. 2003. Is there a 'strength effect' in automatic semantic priming? *Memory and cognition* 31:262–72.

Anderson, John. 2000. *Cognitive psychology and its implications*. New York: Worth.

Aslin, Richard N., Jenny R. Saffran & Elissa L. Newport. 1998. Computation of conditional probability statistics by 8-month-old infants. *Psychological science* 9:321–24.

Barnes, J. M., & B. J. Underwood. 1959. 'Fate' of first-list associations in transfer theory. *Journal of experimental psychology* 58:97–105.

Barry, C., K. W. Hirsh, R. A. Johnston & C. L. Williams. 2001. Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *Journal of memory and language* 44:350-75.

Baysari, Melissa T. & Robert A. Boakes. 2004. Flavour aversion produced by running and attenuated by prior exposure to wheels. *Quarterly journal of experimental psychology* 57B:273–86.

Bonin, Patrick, Michel Fayol & Marylène Chalard. 2001. Age of acquisition and word frequency in written picture naming. *Quarterly journal of experimental psychology* 54A:469–89.

Bowers, Jeffrey S. 2000. In defense of abstractionist theories of repetition priming and word identification. *Psychonomic bulletin and review* 7:83–99.

——— & Chad J. Marsolek (eds.) 2003. *Rethinking implicit memory.* New York: Oxford University Press.

Brown, G. D. & F. L. Watson. 1987. First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory and cognition* 15:208-16.

Castles, Anne, Chris Davis & Tessa Letcher. 1999. Neighbourhood effects on masked form priming in developing readers. *Language and cognitive processes* 14:201–24.

Chwilla, Dorothee J., Peter Hagoort & C. M. Brown. 1998. The mechanism underlying backward priming in a lexical decision task: Spreading activation versus semantic matching. *Quarterly journal of experimental psychology* 51A:531–60.

Coltheart, Max, Eileen Davelaar, Jon Torfi Jonasson & Derek Besner. 1977. Access to the internal lexicon. In *Attention and performance VI*, ed. by Stalislav Dornič, 535–55. Hillsdale NJ: Erlbaum.

Emmorey, Karen, David Corina & Ursula Bellugi. 1995. Differential processing of topographic and referential functions of space. In *Language, gesture and space*, ed. by Karen Emmorey & Judy S. Reilly, 43–62. Hillsdale NJ: Erlbaum.

Feldman, Laurie Beth. 2003. What the repetition priming methodology reveals about morphological aspects of word recognition. In *Rethinking implicit memory*. In Bowers & Marsolek, 124–38.

Forster, K I. & C. Davis. 1984. Repetition priming and frequency attenuation in lexical access. *Journal of experimental psychology: Learning, Memory, and Cognition* 10:680–98.

Ghyselinck, Mandy, Michael B. Lewis & Marc Brysbaert. 2004. Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta psychologica* 115:43–67.

Glanzer, Murray & S. L. Ehrenreich. 1979. Structure and search of the internal lexicon. *Journal of verbal learning and verbal behavior* 18:381–98.

Goldinger, Stephen D., Paul A. Luce & David B. Pisoni. 1989. Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of memory and language*, 28:501–18.

Hall, Geoffrey 2003. Learned changes in the sensitivity of stimulus representations: Associative and nonassociative mechanisms. *The quarterly journal of experimental psychology* 56B:43–55.

Hintzman, Douglas L. 1986. 'Schema abstraction' in a multiple-trace memory model. *Psychological review* 93(4):411–28.

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In *Talker variability in speech processing*, ed. by Keith Johnson & John W. Mullennix, 145–65. San Diego: Academic Press.

Joordens, Steve & Suzanna Becker. 1997. The long and short of semantic priming effects in lexical decision. *Journal of experimental psychology: Learning, Memory, and Cognition* 23:1083–1105.

Kamin, L. J. 1969. Predictability, surprise, attention and conditioning. In *Punishment and aversive behavior*, ed. by B. A. Campbell & R. M. Church, 279–98. New York: Appleton-Century-Crofts.

Kapatsinski, Vsevolod M. 2005. Frequency, age-of-acquisition, lexicon size, neighborhood density and speed of processing: Towards a domain-general, single-mechanism account. *Proceedings of HDLS VI.*

Koriat, A. 1981. Semantic facilitation in lexical decision as a function of prime-target association. *Memory and cognition* 9:587–98.

Luce, Paul A., Stephen D. Goldinger, Edward T. Auer & Michael S. Vitevitch. 2000. Phonetic priming, neighborhood activation, and PARSYN. *Perception and psychophysics* 62:615–25.

Marchant, H. G. & J. W. Moore. 1973. Blocking of the rabbit's conditioned nictitating membrane response in Kamin's two-stage paradigm. *Journal of experimental psychology* 101:155-8.

McGovern, J. B. 1964. Extinction of associations in four transfer paradigms. *Psychological monographs* 78.

Moder, Carol L. 1992. *Productivity and categorization in morphological classes*. Ph.D. diss.: SUNY Buffalo.

Monsell, Stephen. 1991. The nature and locus of word frequency effects in reading. In *Basic processes in reading: Visual word recognition*, ed. by Derek Besner & Glyn W. Humphreys, 148–97. Hillsdale NJ: Erlbaum.

Morton, J. 1969. Interaction of information in word recognition. *Psychological review* 76:165–76.

Nation, Kate & Margaret J. Snowling. 1998. Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child development* 69:996–1011.

Neely, James H. 1991. Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic processes in reading: Visual word recognition*, ed. by Derek Besner & Glyn W. Humphreys, 264–336. Hillsdale NJ: Erlbaum.

Norris, D. G. 1984. The effects of frequency, repetition and stimulus quality in visual word recognition. *Quarterly journal of experimental psychology* 36A:507–18.

Paap, Kenneth R., James E. McDonald, Roger W. Schvaneveldt & Ronald W. Noel. 1987. Frequency and pronounceability in visually presented naming and lexical decision tasks. In *Attention and performance XII*, ed. by Max Coltheart, 221–44. Hove: Erlbaum.

Perea, Manuel & E. Rosa. 2000. Repetition and form priming interact with neighborhood density at a short stimulus-onset asynchrony. *Psychonomic bulletin and review* 7:668–77.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In *Frequency and the emergence of linguistic structure*, ed. by Joan Bybee & Paul J. Hopper, 137–57. Amsterdam: Benjamins.

Plaut, David C. & James R. Booth. 2000. Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological review* 107:786–823.

Radeau, Monique, José Morais & Juan Segui. 1995. Phonological priming between monosyllabic spoken words. *Journal of experimental psychology: Human perception and performance* 21:1297–1311.

Ratcliff, Roger & Gail McKoon. 1981. Does activation really spread? *Psychological review* 88:454–62.

——— & ———. 1988. A retrieval theory of priming in memory. *Psychological review* 95:385–408.

Rueckl, Jay G. 2003. A connectionist perspective on repetition priming. In Bowers & Marsolek, 67–104.

Scarborough, D. L., C. Cortese & H. S. Scarborough. 1977. Frequency and repetition effects in lexical memory. *Journal of experimental psychology: Human perception and performance* 3:1–17.

Schriefers, H., A. Frederici & P. Graetz. 1992. Inflectional and derivational morphology in the mental lexicon: Symmetries and asymmetries in repetition priming. *Quarterly Journal of Experimental Psychology* 44A:373–90.

SCHUBERTH, R. E. & P. EIMAS. 1977. Effects of context on the classification of words and non-words. *Journal of experimental psychology: Human perception and performance* 3:27–36.

SIKSTROM, S. 2002. Forgetting curves: Implications for connectionist models. *Cognitive Psychology* 45:95–152.

SIMPSON, GREG B. & THOMAS C. LORSBACH. 1983. The development of automatic and conscious component of contextual facilitation. *Child development* 54:760–72.

STARK, CRAIG EDWARD LELAND. 1997. *Repetition priming of words, pseudowords and nonwords.* Ph.D. diss.: Carnegie Mellon University.

STEYVERS, MARK & JOSHUA B. TENENBAUM. 2005. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive science* 29:41–78.

STOCKALL, LINNEA. 2004. *Magnetoencephalographic investigations of morphological identity and irregularity*. Ph.D. diss.: MIT.

SUMIDA, RONALD ALAN. 1997. *Parallel distributed semantic networks for natural language processing*. Ph.D. diss: UCLA.

——— & MICHAEL G. DYER. 1992. Propagation filters in PDS networks for sequencing and ambiguity resolution. *Neural information processing systems* 4:233–40.

THOMSEN, C. J., H. LAVINE & J. KOUNIOS. 1996. Social value and attitude concepts in semantic memory: Relational structure, concept strength and the fan effect. *Social cognition* 14:191–225.

VERSACE, R. & B. NEVERS. 2003. Word frequency effect on repetition priming as a function of prime duration and delay between prime and target. *British journal of psychology* 94:389–408.

ZEELENBERG, RENÉ, DIANE PECHER & JEROEN G. W. RAAIJMAKERS. 2003. Associative repetition priming: A selective review and theoretical implications. In Bowers & Marsolek, 261–83.