Units of representation in visual word recognition

Matthew H. Davis*

Medical Research Council Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom

As you read these words, a complex sequence of processes are at work in your brain, identifying visual patterns (letters) that are mapped onto familiar units (words), the meanings of which are combined to allow comprehension. In this description, a mental dictionary or lexicon linking word forms (orthography) to word meanings (semantics) plays a central and critical role in the reading process. However, some fundamental questions concerning the functional and neural organization of the mental lexicon remain unanswered. How are words composed of more than one unit (such as *darkness*) stored in the lexicon? Are they stored as whole forms, or are complex words broken down into their constituent parts or morphemes (*dark* and *ness*)? In this issue of PNAS, Devlin *et al.* (1) describe a functional MRI (fMRI) study that makes significant progress in understanding the functional and neural architecture of the systems involved in accessing the meaning of written words, and provide evidence in support of a controversial approach to the processing of complex words.

Morphemes in the Mental Lexicon

Traditional linguistic accounts of how complex words are stored propose that the mental lexicon is organized morphemically. Morphemic organization ensures that there is no redundancy in the representation of related words created by using either derivational (e.g., *trusty*, *distrust*, *untrustworthy*) or inflectional (*jumps*, *jumped*, *jumping*) morphemes. These theoretical arguments have been supported by the results of psycholinguistic experiments. For instance, responses to a simple word (*hunt*) are speeded or primed by a prior presentation of a related word (*hunter*), suggesting that these words have shared entries in the mental lexicon (2). However, priming could also arise from shared meaning or shared orthographic form, and research has also demonstrated that priming effects for morphologically related word pairs (*hunter–hunt*) are distinct from priming effects for items that have an equivalent amount of overlap in form (*planet–plan*) or in meaning alone (*imitate–copy*) (3). On the basis of these and other findings, models of word recognition typically include a processing stage in which complex words are split into their constituent morphemes before
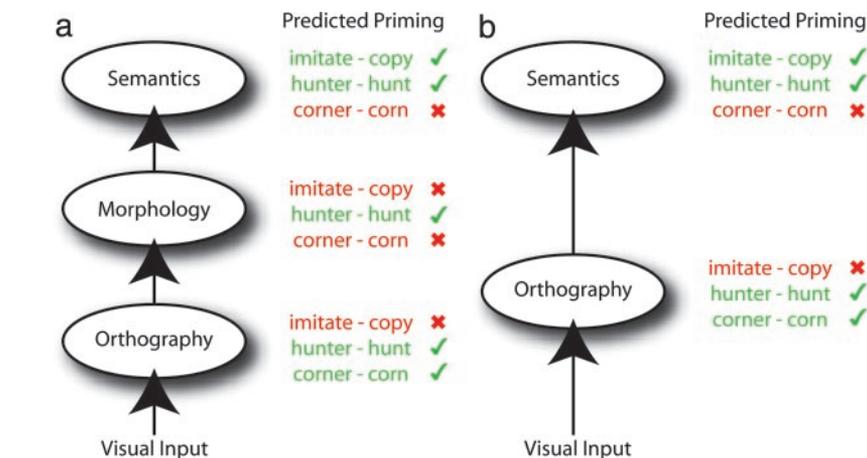

Fig. 1. Two contrasting accounts of the processing stages involved in accessing the meaning of written words and their predictions for whether neural priming effects at different levels will be present (✓) or absent (✗). (*a*) Traditional accounts of word recognition incorporate a ''morphological'' level. Neural priming at this level would be greater for morphologically related items (*hunter–hunt*) than for items with only orthographic or semantic overlap. (*b*) A connectionist account of word recognition suggested by the results of Devlin *et al.* (1) in which there is no morphological level and neural priming effects for morphologically related words overlap with priming for orthographically and semantically related pairs.

meaning-based representations are accessed (4) (Fig. 1*a*).

Despite general agreement that some morphologically complex words have a lexical entry that is shared with related forms, there is disagreement concerning precisely which complex words are stored as morphemic units. For instance, evidence of decomposition into morphemes is typically stronger for words formed with inflectional endings than those formed with derivational endings (2), words that have a clear semantic relationship with their stem (5), or words with morphological endings that are used to form novel words (i.e., productive morphemes; ref. 6). Evidence of varying degrees of decomposition have resulted in more elaborate accounts that propose multiple processing pathways between orthographic and semantic representations (5, 6).

Distributed Connectionist Accounts

Partly in response to evidence for graded decomposition, an alternative account based around distributed connectionist models has been proposed (7, 8). In these connectionist accounts, words and morphemes are not discrete units, but instead patterns of activity over simpler units representing domains of knowledge such as orthography and semantics. Words are stored as the strength of connection weights linking these various units. Critically, a single

set of connections serves double duty by encoding the morphemes of complex words that both can and cannot be decomposed, without requiring a specifically "morphological" level (Fig. 1*b*).

In this description, mapping from word form to word meaning is another linguistic domain that plays host to a debate between dual-route, localist, and single-route, connectionist accounts of cognition (9, 10). Similar debates have endured in the literature on reading aloud (11, 12) and regular and irregular inflection (13, 14). Methods from neuropsychology and cognitive neuroscience have brought novel sources of data to these debates and led to further refinements and greater theoretical specificity (15, 16). There is an increasing acceptance that theories of the acquisition, skilled performance, and breakdown of linguistic processes must be grounded in underlying neural mechanisms if they are to be of value. It is in the brain that the "fact of the matter" lies.

Masked fMRI Priming

Devlin *et al.* (1) make an important contribution to our understanding of the neural systems involved in recognizing complex words. Whereas behavioral

See companion article on page 14984.

*E-mail: matt.davis@mrc-cbu.cam.ac.uk.

© 2004 by The National Academy of Sciences of the USA

priming produces only a single measure (typically response time), priming-related reductions in the magnitude of the fMRI signal can be observed in multiple brain areas, each concerned with processing different aspects of a stimulus (17). For instance, fMRI studies by Dehaene and colleagues (18, 19) have used masked priming to characterize a sequence of visual areas in the fusiform gyrus that generate an abstract representation of printed words, independent of the case and retinal position of the constituent letters. Therefore, neural priming can reveal dissociable processing stages where behavioral priming might reflect the operation of only a single component of this system.

Devlin *et al.* (1) used this masked fMRI priming method to probe neural systems that are central to visual word recognition, testing priming effects for orthographically (*corner–corn*) and semantically (*imitate–copy*) related word pairs, as well as morphologically related pairs that overlap in both orthography and semantics (*hunter–hunt*). Orthographically related word pairs showed priming in the fusiform gyrus, consistent with a role in form processing (18). Priming effects for semantically related pairs were observed in the posterior middle-temporal gyrus, a region showing semantic priming in other studies (20). However, neural priming for morphologically related pairs entirely overlapped with orthographic and semantic effects, despite retrospective power analyses demonstrating adequate sensitivity to detect priming effects.

Implications for Visual Word Recognition

Two aspects of this study are especially noteworthy and illustrate the potential for brain imaging to spur further refinements in accounts of visual word recognition. One feature of Fig. 1*b* that will be surprising to connoisseurs of distributed connectionist accounts is the absence of "hidden units" that mediate between orthographic and semantic rep-

resentations. Connectionist models use these units to encode regularities in meaning that are provided by morphological structure. These hidden units therefore develop overlapping representations for consistent form-meaning pairings such as *hunter–hunt*, although not for opaque pairs like *corner–corn* (7). Indeed, neural network simulations predict less behavioral priming for opaque than for transparent forms (8).

Neural priming for morphologically related pairs entirely overlapped with orthographic and semantic effects.

If the neural priming effects observed by Devlin *et al.* rule out a localist account with explicit, morphological units for decomposed items, then they might similarly rule out a distributed account in which similar representations emerge from the operation of a learning algorithm.

This conclusion might appear to challenge existing distributed connectionist accounts of morphological processing, which have used hidden units (7, 8). Perhaps this discrepancy merely reflects the distance that exists between implemented models and the neural systems that these computational accounts seek to mimic. Neuroanatomy tends to favor local connections within cortical areas. So, rather than hidden units being located in a separate cortical area "between" orthography and semantics, they may instead correspond to neurons and connections within regions that separately represent orthographic and semantic knowledge. Connectionist accounts that incorporate local hidden units have been proposed for other do-

mains (21); further simulations will be required if we are to know whether this change in neural organization has functional consequences for models of morphological processing.

A second aspect of Devlin *et al.*'s results that parallels recent developments in accounts of visual word recognition is the absence of any difference between neural priming for items with (*hunter–hunt*) and without (*corner–corn*) semantic overlap. Recent behavioral research using masked priming has similarly shown an equivalent amount of priming for these pairs, although items like *brothel–broth*, which have orthographic overlap without an affix-like ending (because -*el* is not a morpheme) do not show priming (22). Because neither *corner–corn* nor *brothel–broth* are semantically related, these findings have motivated a form of morphological decomposition that operates solely at an orthographic level, perhaps implemented by local connectivity within the orthographic system. If pairs such as *brothel–broth* were tested in fMRI, they may not show orthographic priming in anterior regions of the fusiform gyrus (forward from $y = -60$; see ref. 19). This result would imply that the orthographic representations encoded in this region are informed by morphological structure even if they do not correspond to the minimal, meaning-bearing elements proposed by classical theories.

This discussion makes clear that morphology will remain useful for describing the forms of linguistic knowledge encoded in particular brain regions, even if morphology does not stand alone as an independent level of neural representation. The results presented by Devlin *et al.* (1) demonstrate one extremely effective technique for probing the internal operations of neural systems critical for word recognition and are sure to inspire future experimental and theoretical developments.

1. Devlin, J. T., Jamison, H. L., Matthews, P. M. & Gonnerman, L. M. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 14984–14988.
2. Stanners, R. F., Neiser, J. J., Hernon, W. P. & Hall, R. (1979) *J. Verbal Learn. Verbal Behav.* **18,** 399–412.
3. Napps, S. E. (1989) *Mem. Cognit.* **17,** 729–739.
4. Giraudo, H. & Grainger, J. (2000) *Lang. Cognit. Process.* **15,** 421–444.
5. Marslen-Wilson, W., Tyler, L. K., Waksler, R. & Older, L. (1994) *Psychol. Rev.* **101,** 3–33.
6. Bertram, R., Schreuder, R. & Baayen, R. H. (2000) *J. Exp. Psychol. Learn. Mem. Cognit.* **26,** 489–511.
7. Rueckl, J. G. & Raveh, M. (1999) *Brain Lang.* **68,** 110–117.

8. Plaut, D. C. & Gonnerman, L. M. (2000) *Lang. Cognit. Process.* **15,** 445–485.
9. Page, M. (2000) *Behav. Brain. Sci.* **23,** 443–512.
10. Elman, J., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking Innateness* (MIT Press, Cambridge, MA).
11. Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001) *Psychol. Rev.* **108,** 204–256.
12. Plaut, D. C., McClelland, J. L., Seidenberg, M. S. & Patterson, K. (1996) *Psychol. Rev.* **103,** 56–115.
13. Joanisse, M. F. & Seidenberg, M. S. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 7592–7597.
14. Pinker, S. (1999) *Words and Rules: The Ingredients of Language* (Weidenfeld & Nicholson, London).
15. Marslen-Wilson, W. & Tyler, L. K. (1998) *Trends Cognit. Sci.* **2,** 428–435.

16. McClelland, J. L. & Patterson, K. (2002) *Trends Cognit. Sci.* **6,** 465–472.
17. Naccache, L. & Dehaene, S. (2001) *Cereb. Cortex* **11,** 966–974.
18. Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B. & Rivière, D. (2001) *Nat. Neurosci.* **4,** 752–758.
19. Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D. & Cohen, L. (2004) *Psychol. Sci.* **15,** 307–313.
20. Giesbrecht, B., Camblin, C. C. & Swaab, T. Y. (2004) *Cereb. Cortex* **14,** 521–529.
21. Plaut, D. C. (2002) *Cognit. Neuropsychol.* **19,** 603–639.
22. Rastle, K., Davis, M. H. & New, B. (2004) *Psychonomic Bull. Rev.*, in press.