Visual word recognition / Morphological processing

9.59 / 24.905
April 21, 2005

Ted Gibson

A brief taxonomy of writing systems

Writing systems: human inventions

All represent words at some level.

- word/morpheme based: **logographic**. e.g., Chinese (morpheme-based)
- syllable-based: **syllabic**. e.g., Japanese kana system. Only about 100 distinct syllables in Japanese, so this is possible. (about 1000 distinct syllables in English).
- phoneme-based (approximately): alphabetic. e.g., English, Spanish, Russian

  Chinese: tonal language. Therefore, a syllable- or phoneme-based writing system would be difficult. 4500 characters to represent around 45,000 words. (But not pure logography: many cues to pronunciation)

Alphabetic languages

Most monotonal languages like English, Spanish are alphabetic (few are syllabic).

Alphabetic languages vary widely in terms of the correspondence between phonemes and letters:

- Shallow orthography: Finnish, Spanish, Serbo-Croation
- Deep orthography: English, Hebrew: The same letter is often used to signify different sounds in different contexts.

Slight trade-off between deep and shallow orthography: deep orthographic languages spell morphologically related but different sounding words similarly. E.g., sign/signal/signify

English: non-optimal: spelling rules are influenced from two very different language families: Germanic and Romance. So English spelling is often seemingly arbitrary and exception-ridden.

Word perception: Some initial observations

Task: Find all the C's:

GQSJCGQRPRCDBCRCPDGGPQCRCD

GQSJCGQRPRCDBCRCPDGGPQCRCD

XCLNZCFNVAKZVXCLNIKWYZXCLKN

XCLNZCFNVAKZVXCLNIKWYZXCLKN

Task: Find all the S's:

MZSTYLKVSHWLVZSTXNXWFKYSFNT

MZ**S**TYLKV**S**HWLVZ**S**TXNXWFKY**S**FNT

PQCPSGOCSRQPDJSDQRBJRPSGQOP

PQCP**S**GOC**S**RQPDJ**S**DQRBJRP**S**GQOP

Word perception: Some initial observations

In-class demonstration: recognizing symbols based on their component features.

Result: It is hard to find a specific letter when it is hidden among letters with similar features.

Word perception:

- Do we identify the letters and then the words?
- Or do we read the words "holistically", relying on their shapes?
- Or do we somehow exploit our knowledge of what letters are in what words to see both letters and words more readily?

Word perception

The evidence suggests that the process is not a uni-directional pass from letter-features to letters to words:

1. Task: Present strings of letters and ask for report of letters: GREAT easier than TRAGE, which is easier than ARGTE

2. Reporting the fifth letter in GREAT is easier than reporting the single letter in -----T

The Word Superiority Effect (Wheeler, Reicher)

But is it actually easier to see the letters in a word context? Or can they just be guessed more easily even if you miss one of the letters?

Task: Stimulus → Mask → Forced choice: Which did you see, K or D?

Condition 1

---K

####

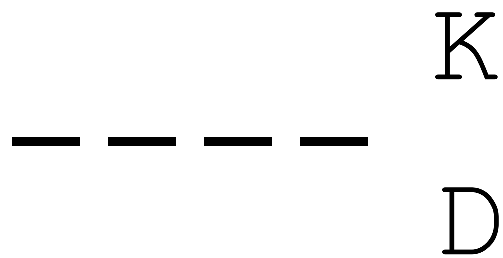
K

D

Condition 2

WORK

####


K
D

The Word Superiority Effect (Wheeler, Reicher)

Participants didn't know which letter in the word would be tested.

Important note: Both --D and --K make possible words.

Result: 10% improvement with the whole word relative to single letter.

Conclusion: It is easier to identify a letter in the context of a word than in isolation.

Addition result: Some benefit for pronounceable non-words like REET, MAVE, WORP.

A parallel distributed model of visual word recognition: McClelland & Rumelhart, 1981

Possible words constrain what letters are seen. Tentative letter assignments constrain each other.

Diagrams removed for copyright reasons.

Several slides removed for copyright reasons.
Figures from McClelland & Rumelhart, 1981

Two routes to visual word recognition

- McClelland & Rumelhart's model: the **direct route** from the visual word to the lexicon.
- An alternative: the **assembled route** hypothesis: Bottom-up visual analysis that is used to activate the phonemes, which when put together, activate the word.
- A third possibility (probably correct): **dual-route**: both the direct and assembled paths are used.

Cross-linguistic evidence of two different routes: Frost, Katz & Bentin (1987)

Priming evidence from three languages:

- Hebrew: deep orthography
- English: quite deep orthography, but less so than Hebrew
- Serbo-Croatian: shallow orthography: one-to-one correspondence between letters and phonemes

Cross-linguistic evidence of two different routes: Frost, Katz & Bentin (1987)

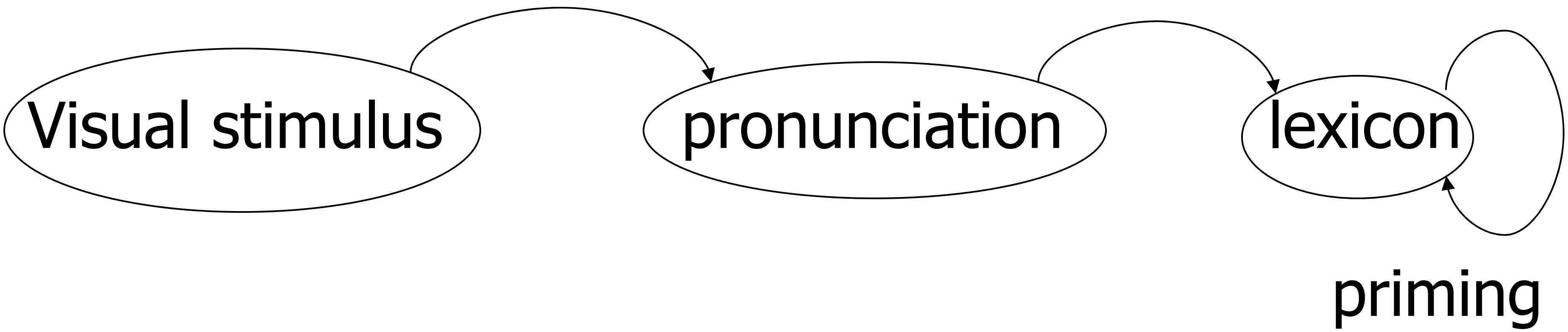
Task: Name target words as fast as possible, preceded by possible prime words.

If using a direct route, then the pronunciation of the word comes after lexical access.

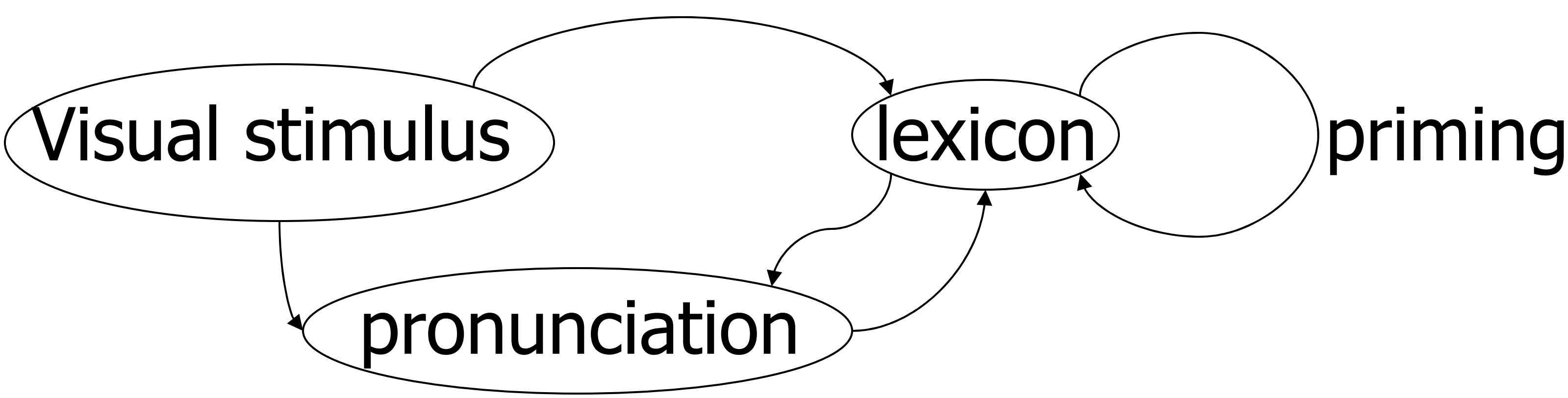
If using an assembled route, then the pronunciation of the word comes before lexical access.

Cross-linguistic evidence of two different routes: Frost, Katz & Bentin (1987)

Shallow orthography



Deep orthography



Cross-linguistic evidence of two different routes: Frost, Katz & Bentin (1987)

Result: Semantic priming in Hebrew and English (less so in English), none in Serbo-Croatian.

This suggests that Hebrew readers rely on a direct route, Serbo-Croatian readers rely on an assembled route, and English speakers may rely on both.

Evidence for assembled route in English: Van Orden (1987): In-class demonstration

- Task: Do the following words indicate animate things? (Animals, people, occupations, etc.)
  - ➢ Note: the words might be anything:

    DOG → "Yes"

    PUSH → "No"

    OFF → "No"

PRINCE

NONE

SELLER

PROPHET

FLEE

FOWL

TOWED

WAIL

PIGEON

NIGHT

PRINCE: yes

NONE: no (control: NOUN)

SELLER: yes

PROPHET: yes

FLEE: no (control: FLEX)

FOWL: yes

TOWED: no (control: TROD)

WAIL: no (control: WHACK)

PIGEON: yes

NIGHT: no (control: KNIFE)

Evidence for assembled route in English: Van Orden (1987): In-class demonstration

Result: People make more errors in **rejecting** targets that do not belong to a category if the target is a homophone of the member of the category.

E.g., harder to say “no” to animate – “none”
than to animate – “noun”

or: food – “meet” vs. food – “melt”

(Note: length, frequency, and visual similarity to the homophone are controlled)

Evidence for assembled route in English: Van Orden (1987): In-class demonstration

Jared & Seidenberg (1991): The homophone effect occurs mainly for low-frequency words.

Therefore, the meaning of low-frequency words may be accessed by an assembled route, while the meaning of high-frequency words may be accessed by a direct route.

Sublexical processing: Syllables

Syllabic processing evidence (Prinzmetal, Treiman & Rho 1986):

A participant was shown a target letter (e.g., “d”), then a word containing the target letter. The word was shown in two colors. The participant's task was to say what color the letter was:

target: d
word: VODKA

Illusory conjunction effect: mistakenly say that one letter is the color of one of its neighbors.

Result: many more illusory conjunction effects when the color boundary was inconsistent with the syllable boundary.

E.g., VODKA : “d” likely to be misclassified.
VODKA : “d” not likely to be misclassified.

WARNING: There are colors on this slide, and the colors are crucial to understanding the results!

Sublexical processing: Syllables

Seidenberg (1987) observation: the original experiment didn't control for **bigram frequency**: pairs of letters. e.g., “dk” much less frequent than “od” or “ka”.

So the result is consistent with the hypothesis that people are classifying according to frequent bigrams.

Rapp (1992) controlled for bigram frequency, and the syllable effect remains.

Sublexical processing: Morphemes

- “jump" primes "jump" in lexical decision (no surprise there)
- "jumped" primes "jump" just as much, suggesting that "jump" is accessed.
- Much less priming in "select" after seeing "selective".

Generally, little priming in derivational morphology (category/ unpredictable meaning changing morphology), but lots in inflectional morphology.

In derivational morphology, when the category is unchanged, and the meaning is compositional, then there is more priming:

"dishonest" primes "honest” but "arson" does not prime "son".

Sublexical processing: Morphemes

Seidenberg (1987): Perhaps all a related meaning effect, not morphology at all.

Evidence against this: Segment-shifting task (Feldman, 1994): A participant is shown a word along with an indication of a segment of that word which is to be added to another word:

HARDEN then BRIGHT

Response: BRIGHTEN

Participants are faster when the segment is a morpheme of the word (as above) compared with controls like:

GARDEN then BRIGHT

Conclusion: People can break up words more easily at morpheme boundaries. This doesn't depend on semantic overlap.

Frequency effects

High frequency words are recognized faster than low frequency words (!)

More interestingly, the direct route seems more important in high frequency words than in low frequency words (Jared & Seidenberg, 1991):

Low-frequency words that follow regular spelling-to-sound correspondence rules are recognized (named, or lexical decision) more quickly than low-frequency words that don't follow regular spelling-to-sound correspondence rules.

E.g., PLUMP is recognized faster than CASTE

But no difference among high-frequency words.

E.g., STOP is recognized at the same speed as SAYS (pronounced sez)

This is called a frequency-by-regularity interaction.

Frequency effects

Modeling these results: Dual route processing.

A race between the two routes:

For high-frequency words, the direct route usually wins. For low-frequency words, the assembled route usually wins.

Neighborhood effects

Neighborhood effects: The influence of lexical similarity in appearance.

Neighborhood: the number of words that can be formed by changing one letter.

E.g., two neighbors of "kind" are "king" and "find".

Andrews (1989, 1992): For low frequency words, recognition is faster for words from large neighborhoods.

This result is predicted by the interactive activation approach, but not by a model proposing an item-by-item search of the lexicon (more like what you do when looking up a word in a dictionary).

However, contrary to the interactive activation approach, there is no neighborhood effect for high-frequency words. This perhaps reflects a floor effect in the top-down processing: the direct access is too quick to get any help from the top-down influence.

Context effects

Question: Does the context speed word recognition?

Priming effects suggest that the answer is yes, but there is an alternative interpretation of priming effects:

Perhaps the word recognition process is the same, but a later stage is affected: e.g., in a lexical decision task, decision-making about whether we have identified the right word.

Context effects

Word recognition without a lexical decision task: In normal reading.

Zola (1984): Lexical priming contexts in full sentence reading (eye-tracking):

(1) Movie theaters must have buttered popcorn to serve their patrons.

(2) Movie theaters must have adequate popcorn to serve their patrons.

Look for speeded RTs on “popcorn” in (1) relative to (2): Small but reliable effect: 15 msec.

Context effects

More developed contexts (Morris, 1994):

(1) The friend talked as the person trimmed the mustache after lunch.
(2) The friend talked to the barber and trimmed the mustache after lunch.
(3) The friend talked as the barber trimmed the mustache after lunch.

Big priming effects on “mustache” in (3) relative to (1). A small lexical priming effect in (2) relative to (1) (as in Zola’s results above).

Developmental dyslexia

Dyslexia or **Specific reading disability**: no obvious cause for someone's inability to master reading.

Not general intellectual impairment, or poor vision, or a lack or opportunity or low motivation.

When these have been ruled out, then dyslexia is diagnosed by default.

About 4% of the population are diagnosed with developmental dyslexia (as opposed to **acquired dyslexia**, which is caused by a brain injury later in life).

Can have devastating consequences: low self-esteem, employability

A view common in the media is that dyslexia is a subtle problem of visual processing, such that visual information sometimes gets "turned around", resulting in seeing or producing reversed letters in writing.

This is not actually a good cue for dyslexia.

Developmental dyslexia

What we have learned from normal visual word recognition: dual route.

There are two primary kinds of dyslexics: **phonological** dyslexics and **surface** dyslexics (some are also mixed: both problems).

Developmental dyslexia

Phonological dyslexics make up about 60% of dyslexics: Have trouble with the assembled route. Normal reading on high-frequency words, but very poor on low frequency words that need to be sounded out.

E.g., case study of an adolescent boy J.M. (Snowling et al., 1994):

Reading at about 5 years below his age level: 1) Normal for his age group on high frequency words; 2) Much worse than normal for lower frequency words, even worse than normal for a child five years younger.

Developmental dyslexia

Surface dyslexics: sound out all words laboriously, like beginning readers. Some problem with visual perception or visual memory has prevented them from developing good bottom-up visual processing.

Treatment for dyslexia? Different treatments for different impairments. For many, intensive training in spelling-to-sound rules.

Visual word recognition: Conclusions

There are two routes to word recognition:

The **direct** route (for high frequency items) and the **assembled** route (for low frequency items).

Morphological processing

What is the primary unit of storage in the mental lexicon? Words? Morphemes?

morpheme: the smallest meaning-bearing unit in the language

happy: one morpheme: a stem, {happy}
unhappily: {un} + {happy} + {ly}
happiness: {happy} + {ness}
running: {run} + {ing}

Morphological processing

Words = stems and affixes (prefixes, suffixes, infixes)

Two kinds of affixes: **Inflectional** and **Derivational**

Inflectional affixes: for grammatical agreement in a sentence. The syntactic category of the word remains unchanged. Role-assignment remains unchanged.

English nouns: 2 forms: duck, ducks

English verbs: 4 forms: quack, quacks, quacking, quacked

Italian, Spanish: 50 verb forms

classical Greek: 300-400

Turkish: 2 million

Morphological processing

Derivational affixes: Often change the syntactic category and thematic-role assignment properties of a stem / word. E.g. N to V, V to Adj, etc.

-able: changes a Verb "to do X" into an Adjective: "capable of having X done to it"

learn-learnable, teach-teachable, hug-huggable

Morphological processing

What is the primary unit of storage in the mental lexicon?

Are morphologically complex words represented as full forms, or does the representation reflect their morphological structure?

The full-listing hypothesis

vs.

The morphemic (decompositional) hypothesis

Morphological processing

A variety of possible models, depending on full-listing/decomposition in representations and access:

- Full-listing of forms, full-listing of meanings (Bybee, 1988). Thus, there are fully independent representations of "happy" and "happiness", and "run" and "running‟.

- Full-listing access by forms, decomposed representations;

- Decomposed access, decomposed representations (Marslen-Wilson & Tyler, 1998; Allen & Badecker, 1999);

- Intermediate models: sometimes decomposed access, sometimes full-listing access (Caramazza et al., 1988; Schreuder & Baayen, 1995).

Morphological processing

- A strong full-listing account is not tenable crosslinguistically: languages like Turkish and Finnish are highly productive morphologically, with over one million forms of each verb. It is not plausible that these are all stored independently.

Morphological processing: Potential evidence for form-level decomposition

- Compare response times, error rates or eye-fixation times for sets of complex words: more frequent items are recognized faster, with lower error rates.

  Taft (1979) result: Whole word frequency can be controlled, and there can still be a difference in RTs, if the stems differ in frequency

  E.g., "sized" has the same frequency as "raked", but "size" is more frequent than "rake":

  "sized" is recognized faster than "raked"

Morphological processing: Potential evidence for form-level decomposition

Does this mean that stems must exist independently from their affixes?

No: The faster recognition could be due to increased activation due to overlaps in other aspects of the lexical representations (meanings, etc.)

Evidence in support of this alternative interpretation (Kelliher & Henderson, 1990): Irregularly inflected words show the same effects:

"bought" has the same frequency as "shook"; "buy" is more frequent than "shake":

"bought" is recognized faster than "shook"

This difference is not likely due increased activation of the stem's form ("buy"), because "bought" is not composed of "buy" plus an affix.

The observed effects are therefore likely caused by overlaps in the words' meanings: aspects of their lexical representations

Evidence for the decompositional representation of morphologically complex words in English

Priming experiments: Cross-modal lexical decision task: Is this a word or not?

Auditory prime (e.g., “happiness”) followed by visual prime word (e.g., “happy”).

Evidence for the decompositional representation of morphologically complex words in English

Result 1: Words with derivational morphology (e.g., "happiness") prime their stems (e.g., "happy") but only when the relationship between the two is semantically transparent (Marslen-Wilson et al., 1994):

happiness, happy; rebuild, build

but not: apartment, apart; release, lease

Third condition: purely phonological relationship between prime and target:

bulletin, bullet; tinsel, tin

The priming effect is therefore not due to phonological or orthographic overlap in the words.

Graph removed for copyright reasons.

Evidence for the decompositional representation of morphologically complex words in English

But is this evidence for the form of the stem being primed?
No: Alternative interpretation: Semantic priming.

e.g., “happiness” primes “happy" as “doctor” primes “nurse”

Evidence against this claim:

1. Suffixed pairs sharing the same stem do not prime each other:

“excitement” primes “excite” and “excitable” primes “excite”, but “excitement” does not prime “excitable”

Inhibition among derivationally related words.

Evidence against the semantic priming interpretation of morphological priming

2. Different time courses:

At short delays (one intervening item) priming is equally strong for morphologically and semantically related items:

Morphologically related: "excitement" primes "excite" at 39 msec;

Semantically related: "cello" primes "violin" at about 31 msec. (Not significantly different from the morphological priming effect)

At longer delays (eight intervening items) morphological priming is undiminished at 30 msec, but semantic priming has disappeared (1 msec).

Evidence against the semantic priming interpretation of morphological priming

3. Affixes prime each other: No obvious meaning component in the affixes: (but it is not so clear that there isn't a meaning component to the affixes: This is a sketchy argument)

"-ness": "happiness" primes "darkness"

"re-": "refill" primes "rebuild"

These effects interact with productivity: The more productive an affix, the greater the priming effect

Much less priming for "en-" and "-th":

"enslave" does not prime "enlighten"

"depth" does not prime "sixth"

Evidence against the semantic priming interpretation of morphological priming

3. Affixes prime each other: No obvious meaning component in the affixes:

Graph removed for copyright reasons.

Evidence that some complex words are stored at the form level

- Compositional accounts predict base (stem) frequency effects, but no surface frequency effects.

  But for some classes of words, surface frequency effects occur independent of stem frequency. These results suggest that very frequent surface forms are stored at the form level.

Evidence that some complex words are stored at the form level

E.g., Dutch word forms: nouns, verbs (Schreuder & Bayaan, 1995; Baayen et al., 2000):

Dutch perfect participles: formed by adding the circumfix "ge- ... -d"

"grijns" ("grin") → "gegrijnsd" ("grinned")

high surface frequency (18.3 / million words) vs. low surface frequency (0.9 / million words)

matched for base frequency: 118 / million words each;

also matched for family size, and mean length in letters.

Results: Response latencies to lexical decision:

High surface frequency responses were faster (594 msec) than low surface frequency responses (702 msec)

Morphological processing: Conclusions

- In lexical access, words are sometimes decomposed into morphemes, and sometimes (for very frequent words) they are accessed using their full listings.