Seeing Stems Everywhere and Being Blind to Affixes

Positional Constraints on Morpheme Identification

Davide Crepaldi[§], Kathy Rastle[@], Colin J. Davis[@], and Stephen J. Lupker[#]

[§] Department of Psychology, University of Milano-Bicocca, Italy
[@] Department of Psychology, Royal Holloway University of London, UK
[#] Department of Psychology, University of Western Ontario, Canada

davide.crepaldi1@unimib.it

INTRODUCTION

- ✓ Crepaldi, Rastle and Davis (2010) have shown that suffix identification is position-specific in English: suffixes are identified in morphologically-structured nonwords when they follow an existing stem (such as in SHOOTMENT), but not when they come before an existing stem (such as in MENTSHOOT).
- ✓ These results remain unexplained in current theories of visual word identification, as none takes up the issue of morpheme position coding.
- ✓ A simple interpretation of these results is that the morpheme identification system is sensitive to positional constraints, and thus learns that suffixes can only be so if they follow a stem.
- ✓ This hypothesis makes strong predictions on prefixes and free morpheme identification:
  - Prefixes should only be identified in word initial position when followed by a stem (e.g., UNEAR), because they only appear in such configuration in English.
  - Free morphemes should be identified everywhere because they can appear in isolation (free of any positional constraint) and can in principle occupy any position in complex words (e.g., PREHEAT and HEATING, OVERESTIMATE and HANGOVER).

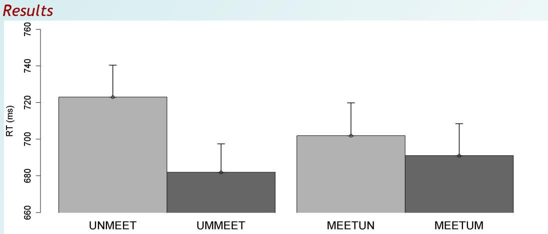
EXPERIMENT 1 - PREFIXES

If prefixes are only identified at word onset, we should observe morpheme interference effects in UNMEET, but not in MEETUN

Methods

- ✓ Lexical decision task designed to elicit morpheme interference effect (e.g., Taft & Forster, 1975)
- ✓ Four nonword conditions:
  - *Prefix-plus-stem*: UNMEET
  - *Control-plus-stem*: UMMEET
  - *Stem-plus-prefix:* MEETUN
  - *Stem-plus-control*: MEETUM
- ✓ Same stems, prefixes and controls across conditions. Nonwords also matched across conditions for ortographic Levenshtein distance to the closest neighbour (OLD1; Yarkoni, Balota & Yap, 2008), mean ortographic Levenshtein distance to the 20 closest neighbours (OLD20; Yarkoni et al., 2008), *N* and mean log bigram frequency.
- ✓ 60 participants, 56 items per condition.
- ✓ 49 simple nonwords, 49 simple words and 56 complex words used as fillers (word/non-word trials = 1; complex/simple targets = 1.14).

Results



- ✓ Interaction between morphological structure and prefix position: $F_1[1, 171] = 5.83$, $p = .01$, $F_2[1, 153] = 1.17$, $p = .28$.
- ✓ Post-hoc analyses show that UNMEET is slower than UMMEET ($t_1[57] = 4.18$, $p <.001$; $t_2[52] = 2.70$, $p < .01$), but MEETUN is not slower than MEETUM ($t_1[57] = 1.328$, $p = .18$; $t_2[51] = .92$, $p = .36$)

EXPERIMENT 2 - FREE STEMS

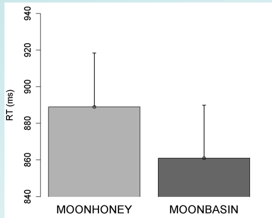
If free morphemes are identified everywhere within words, MOONHONEY should activate HONEYMOON in the lexicon and should thus be hard to reject as a nonword

Methods

- ✓ Lexical decision task designed to elicit morpheme interference effect
- ✓ Two nonword conditions:
  - *Shifted-morphemes compounds*: MOONHONEY
  - *Controls*: MOONBASIN
- ✓ Covariates considered (fully crossed design): position of the substituted morpheme (word-initial vs. word-final), semantic transparency of the word-initial morpheme in the original compound, semantic transparency of the word-final morpheme in the original compound.
- ✓ Related and control morphemes (e.g., HONEY and BASIN) matched for length, written frequency, spoken frequency, *N* and semantic distance to the other morpheme (e.g., MOON) as estimated by Latent Semantic Analysis (Landauer & Dumais, 1997).
- ✓ Non-words matched across conditions for length, *N*, mean log bigram frequency and number of syllables.
- ✓ 48 participants, 48 items per condition.
- ✓ 48 simple nonwords, 48 simple words and 48 complex words used as fillers (word/non-word trials = 1; complex/simple targets = 1).

EXPERIMENT 2 - FREE STEMS (cont'd)

Results



- ✓ MOONHONEY significantly slower than MOONBASIN: $F_1[1, 42] = 7.20$, $p = .01$; $F_2[1, 42] = 7.71$, $p = .008$.
- ✓ Shifted-morphemes effect does not interact with either semantic transparency (in word-initial position: $F_2[1, 42] =.03$, $p = .86$; in word-final position: $F_2[1, 42] =.38$, $p = .55$) or position of the substituted morpheme ($F_2[1, 42] =.40$, $p = .53$).

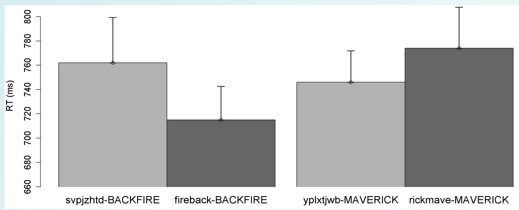
EXPERIMENT 3 - FREE STEMS

If free morphemes are identified everywhere within words, constituent priming should also occur across position

Methods

- ✓ Masked priming study (SOA=48 ms), lexical decision task.
- ✓ 48 compound words and 48 simple words served as targets, each paired with two prime non-words:
  - *Shifted-halves*: fireback-BACKFIRE, rickmave-MAVERICK
  - *Controls*: svpjzhtd-BACKFIRE, yplxtjwb-MAVERICK
- ✓ 2x2 crossed design: Morphological Structure (compounds vs. simple words) and Relatedness (related vs. control primes).
- ✓ Compound and simple word targets matched for length, written frequency, spoken frequency, *N*, mean log bigram frequency, number of phonemes and number of syllables.
- ✓ Related and control primes matched for length, *N*, mean log bigram frequency and OLD1.
- ✓ 24 participants.
- ✓ 96 non-word trials (word/non-word trials = 1; related/unrelated trials = 1).

Results



- ✓ Interaction between morphological structure and relatedness: $F_1[1, 63] = 6.71$, $p = .01$, $F_2[1, 86] = 11.80$, $p < .001$.
- ✓ Post-hoc analyses show that FIREBACK facilitates BACKFIRE ($t_1[21] = 2.45$, $p = .02$; $t_2[46] = 2.91$, $p < .01$), but RICKMAVE does not facilitate MAVERICK (perhaps inhibits, in fact) ($t_1[57] = -1.29$, $p = .20$; $t_2[40] = -1.96$, $p = .06$)

CONCLUSIONS

- ✓ Prefix identification is position-specific
- ✓ Free stem identification is position-independent

Together with the results of Crepaldi et al. (2010), these data show that:

✓ BOUND MORPHEMES are position-locked in the word identification system
✓ FREE MORPHEMES are position-independent in the word identification system

- ✓ Different coding schemes for bound and free morphemes at the morpho-orthographic level?
- ✓ Position needs to be coded further up in the system for free morphemes, or otherwise how could we distinguish between morphological anagrams such as OVERHANG and HANGOVER?
  So, which coding scheme is flexible enough to allow position-independent identification of free stems, but position coding at further levels of representation?

ACKNOWLEDGMENTS

Experiment 1 and 2 were carried out at the Department of Psychology, Royal Holloway University of London, and were supported by the Economic and Social Research Council, the Leverhulm Trust and the British Academy. Experiment 3 was carried out at the Department of Psychology, University of Western Ontario, and was funded by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Crepaldi, D., Rastle, K., & Davis, C.J. (2010). Morphemes in their place: Evidence for position-specific identification of suffixes. Memory and Cognition, 38, 312-321.
Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. Psychological Review, 104, 211-240.
Taft, M. & Forster, K.I. (1975). Lexical storage and retrieval of prefixed words. Journal of Verbal Learning and Verbal Behavior, 14, 638-647.
Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. Psychonomic Bulletin & Review, 15, 971-979.