# First Meeting: Overview of issues and thoughts about Affect ∩ Language

Morgan Moyer
March 19, 2024

---

## Contents

**1 Briefly: A path forward as I see it** **1**

**2 Hypotheses** **2**
2.1 Affect Primacy Hypothesis . . . . . 2
2.1.1 What is affect? . . . . . 2
2.2 Lexical Valence Hypothesis . . . . . 3
2.3 The Emotional Embodied/Grounding Hypothesis . . . . . 4

**3 Experimental Points** **5**
3.1 Experiment 1: Norming / RT . . . . . 5
3.2 Experiment 2: Similarity Judgement . . . . . 5
3.3 Proposed Experiment 3: Priming . . . . . 5
3.4 Another possible study: Corpus/Experimental . . . . . 6
3.5 General Experimental/Methodological points . . . . . 6

**4 Theoretical** **7**
4.1 What is (affective) meaning? . . . . . 7
4.2 Words in sentential context vs. isolation . . . . . 8
4.3 Playing around: Linguistic tests for content types . . . . . 9

---

# 1 Briefly: A path forward as I see it

- Experiments 1 & 2 the same with some small modifications
  - Adding Neutral words
  - Rethinking descriptive foils for valence
- At least adding an Experiment 3 (priming)
  - If valence is lexically represented, then it should give rise to priming effects
  - Lexical decision task

- Potentially Experiment 4 (corpus + experiment), though it may be difficult to execute in 18(16) months
  - capturing the context sensitivity of affective meaning
  - polysemy of event structure / meaning

# 2 Hypotheses

## 2.1 Affect Primacy Hypothesis

- Is affect "post-cognitive" or "pre-cognitive"?
- "It is further possible that we can like something or be afraid of it before we know precisely what it is and perhaps even *without* knowing what it is," (Zajonc 1980, 154)
- "Preference need no inferences"

### 2.1.1 What is affect?

1. ERP/EEG/fMRI studies to break down what exactly is the time course of emotion processing, how many different kinds of emotion/affect response are there
   - Hot vs. cold (Zajonc 1980 and others)
   - Affective vs. Semantic valence (Itkes et al 2017)
   - Emotion laden vs. emotion label words (Pavlenko 2008, Betancourt et al 2024)
2. From ERP studies (with emo words) there are at least two consistent responses:
   - Early posterior negativity (EPN) $\sim$200-300ms post-stimulus occipito-temporal region

     Larger from Emo words than neutral words (e.g., 'erotica' vs. 'neudists' in Farkas, Oliver, Sabatinelli 2019)

     Indicative of lexico-semantic processing; automatic, task-independent attention allocation

   - Late Postive Complex (LPC) $\sim$400ms post stimulus with centro-parietal topography

     Larger for Emo words than Neural words; Sometimes differences found between positive and negative

Evaluation and controlled attention; semantic processing

Modulation by stimulus type and task (=present if task requires lexical processing, absent if structural task)

3. Korber et al (2008) identified 6 functional groups of consistent neural co-activation during emotion processing! Many of these systems involve interactions between them

    - We might try to functionally oppose cortical activity (locus of 'cognition') with $\sim$ brain stem/midbrain ??
    - Brain *circuits*: macro circuits connect distinct brain regions (Malezieux et al 2023)
    - Appetitive vs. aversive responses to sensory stimuli in e.g., taste receptors, but crucially *flexibility* is a neuro(-bio)logical hallmark of emotion (Malezieux et al 2023, Tye 2018, Craige 2003, Liljencrantz & Olausson 2014, Marshall et al. 2019)
    - Biological default dispositions that can be unlearned (amygdala)

4. Is the Affect / Cognition dichotomy still justified?

**How does this hypothesis apply to language?**

[mm: Side thought on lexical items and inferential role semantics, see end]

## 2.2 Lexical Valence Hypothesis

- Valence is encoded in the lexical entry (for some words)
- Q: Just positive/negative? or neutral too?

This hypothesis is introduced at the end of the paper, but it's a linking hypothesis that's needed to bridge emotion and meaning.

**Linking Question**: If affect is lexically represented, in virtue of what would we think that the lexical valence feature would be prior in the way the Affect First Hypothesis would state?

One link is the Emotional Grounding Hypothesis: Since emotion is prior, emotional meaning activaes emotional pathways, therefore is also prior (because affective valence is prior, semantic valence is too)

## 2.3 The Emotional Embodied/Grounding Hypothesis

- Are abstract concepts grounded in emotion(al information)?
- Since abstract concepts lack experiential (=sensori-motor) correlates, perhaps they are grounded in linguistic and in affective information
- There *may* be some evidence for embodiment in lexical processing but it's debated, and small effects (studies cited in Meteyand & Vigliocco 2018)
- Further question, to what extent does emotion ground all concepts?
- Concrete/Abstract vs. Psychological/Physical? See below
- Concreness vs. Valence: the "inverted U" pattern
  - Kousta, Vigliocco, Vinson, Andrews and Del Campo (2011), "The representation of abstract words"
    - Statistically, it seems valence and concreteness form an "inverted u" pattern relationship: more strongly valenced (postive and negative) words are less concrete; less strongly valenced words are more concrete.
    - In contrast to typical result, they found abstract words were processed *faster* than concrete words when the abstract words are highly valenced (pos/neg vs. neutral)
  - Support for: Boots in RT, accuracy, age of acquisition for highly valence abstract words vs. low valence (=neutral) abstract words, ERP effects for abstract but not concrete (Newcombe et al 2012, Pauligk et al, Vigliocco et al, Ponari et al)
  - Support against: highly valenced concrete words processed faster, ERP (EPN) effects only for concrete but not abstract...(Yao et al, Kanske & Kotz, Palazova et al)
  - Winter (2023), "Abstract concepts and emotion: cross-linguistic evidence and arguments against affective embodiment"
    - That statistical tendency, is it real?
    - Well, if you aggregated everything together, usind valence and concreteness as predictors, it seems like yes.
    - However: closer examination of the data reveals several sub-clusters
    - Replacing concreteness with a Cluster predictor explains the data better
    - And actually for several langauges the U-pattern reversed : Dutch, Indonesian, German, French and Mandarin Chinese
    - Abstract concepts are a heterogenous category (Borghi et al 2019; Villani et al 2019; Muraki et al 2020; Desai et al 2018) we need a better measure [mm: focus on sensory experience or perceptual strength]

# 3 Experimental Points

## 3.1 Experiment 1: Norming / RT

- Two Tasks: Affect Judgement Task, Descriptive Task
- Use continuous slider from one to the other. Because in fact, there may be vaguess in the boundaries and we can 'bin' the responses into 3 categories afterwards for model analysis and Exp 2
- Necessary to test **neutral words**, because they are a control for the affect dimension....crucially to validate an emotional response you not only need to show it happens with the emotion word, but that it doesn't happen with the non-emotion word
- RT $\sim$ Task(Affect, Desc)
  Affect: RT $\sim$ Valence (+,-, neutral)
  Desc: RT $\sim$ Conc (+,-)

## 3.2 Experiment 2: Similarity Judgement

- Speeded similarity judgement task.
- Speeded response task logic: by forcing participants to respond as quickly as possible, you're not giving them time to engage in deliberate, effortful reasoning. Thus, so the logic goes, the quickest responses are candidate for automatic cogntiion (usually meaning part of the lexicon)
- Comparison from (L)LMs: GloVE, BERT
- Important to control for salience, b/c valence might be more salient for the similary judgement than the descriptive dimension if the descriptive dimension doesn't accurately capture the verb meanings

## 3.3 Proposed Experiment 3: Priming

- If the Lexical Affect Hypothesis is seriously considered, than a priming study is a good (=well-established) bet.
- Logic: priming occurs when the prime and target words share some kind of representational link (deliberately vague here).
- If valence is a feature represented in the lexical entry for valenced words, then if the prime and target share the feature, priming will occur.
- Priming study
  Condition: Valence (3) x Concrete (2 or 3)
  RT $\sim$ Condition

## 3.4 Another possible study: Corpus/Experimental

- Goal: to capture some of the polysemy in the verb meaning
- Step 1: extract naturalistic utterances of verbs. (+ 10 preceding lines of discourse)
  Step 2: have participants rate those naturalistic occurences for valence, descriptive content.
  Potential Step 3: Compare context effects by doing Step 2 but without the 10 preceding lines of discourse. Compare results.
  (this is a methodology I've done before in my first postdoc with Judith Degen)
- Thus, there might be a high context sensitivity of valence judgements
- It seems that words might change valence in sentential context, i.e., event structure is vague until event participants are specified (think: 'thaw my heart' vs. 'thaw the ice')
- How does the valence of a word change in the naturalistic context of a sentence?
- By quantifying not only the frequency and distribution of valence in naturalistic context, we can compare the ratings from Experiment 1 to this study, matched on item, even as a predictor in the model

## 3.5 General Experimental/Methodological points

1. Control for possible psycholinguistic confounds in the words used:

   1.1. Syntactic/morphological complexity of verb meaning may cause a counfound: transitivity, passivity, tense might affect measure of domain, i.e., instransitive and passive might lead less extreme ratings of physicality
   "The cat pleases me" vs. "I am pleased"
   $\rightarrow$ but these are specified in sentential context

   1.2. Morphological markers of valence might make the valence judgement more salient and therefore faster
   [mm: Response: on the item-by-item basis, are there difference for these morphologically marked verbs compared to the non marked?]

   1.3. Word STD / Variance, how much consensus were there in responses?
   - some words may be more or less valenced, and that may affect RT
   - some words may be more or less polysemous wrt to descriptive domain: "thawing dinner" vs. "thawing emotions"

   1.4. Thoughts about the stimuli (probably related to point 1):
   - Polysemy: Some words differ in psychological/physical dimension in different contexts
   'fabricate a building' vs. 'fabricate evidence'
   'my feelings thawed' vs. 'the ice thawed'

- Passivity? ‘collapse’
  ? ‘I collapsed the building’ vs. ‘The building collapsed’
  ‘that pleases me’ vs. ‘I please you’
- Are ‘disquiet’,‘solace’ really verbs?
- Why is ’kiss’ (2,1) less physical than ‘hug’ (1,8)?

2. Additionally, the literature suggest that the following variables can confound the results (note, some of these may more affect ERP results rather than RT)

   2.1. by-item:

      i. Frequency
      ii. Word length (Hinojoa 2019 et al survey)
      iii. Semantic Association using Word2Vec or BERT (following Souter at al 2023) – Experiment 2
      iv. Influences on abstract/concreteness (Strik Lievers, et al 2021):
         - Lexical category
         - Etymology/morphological structure

   2.2. by-participant:

      i. Gender of part (Warriner et al 2013)
      ii. Native language / billinguallism

3. Notes on general structure, reporting design, materials, methods and results

   - The Lexical Valence Hypothesis isn’t presented until the conclusion
   - Predictions from each hypothesis stated before the experimental results presented

4. Probably a different measure of abstractness, neither abstract/concrete nor physical/psyhological (re: discussion in Winter 2023)

5. Must include Neutral valence words, a critical comparison

# 4 Theoretical

## 4.1 What is (affective) meaning?

**Psycholinguisitcs/cognitive science**
Read: What are lexical items, What are concepts, and how are the two related?

- Local vs. distributed (abstract) representations stored in semantic memory
  - Local: abstract, algebraic
    * Associated network of words
      (Collins & Loftus 1975, Collins & Quillian 1969)

        * Collection of binary features
          (Smith, Shoban & Rips 1974, Tversky 1977)
          Similarity is overlap of features
    - Distributional: co-occurence statistics
      (Harris 1970, Frith 1957, Wittgenstein 1953)

- Format: amodal or embodied ($\sim$ propositional or a-propositional)
  NB: usually 'embodied' is meant to oppose 'information', but given the emerging science of biosemiotics and the gathering evidence for the existence of codes in biology, chemistry and thermodynamics ...

    - If valence, qua semantic/lexical affect, is indeed lexically represented, why would we think it should be accessed faster than descrptive meaning?
    - What is the link between the meaning of lexical affect feature and the primacy of emotion?

**Semantic Theory**

- Referential (Pietroski's "the extension dogma") vs. _____
- The Extension Dogma: "expressions are 'true of' things, and the meaning of expression E determines the things E is true of, "
- What are the candiadate alternatives? Use-based, Pietroski-style, radical pragmatics? something else?
- Pietroski: "Meanings are recipes for how to build mental representations from lexical items...."

**Psycholinguisitcs and the Lexicon**

- If Affect primacy is "preferences need no inferences", then in the case of 'lexical affect' what does this mean? On many theories of the lexicon, components of a lexical entry are inference-generating (Pustejovsky & Jackendoff?, inferential role semantics, tho Fodor & Lepore counterarguments in "Compositionality Papers")
- Can Fodor & Lepore compositionality problems be solved by convergence zones (see Meteyand & Vigliocco 2018 discussion section 4.3.1, citing Damasio 1989, Damasio & Damasio 1994, Marting 2016, Simmons & Barsalou 2003)

## 4.2 Words in sentential context vs. isolation

- If 'love' is lexically positive and 'hate' is lexically negative, what about sign reversals:

    (1) I love racism
    (2) I hate racism

- these show polarity reversals, OR the lexical polarity is neutralised in the sentential context
- alternatively, learning can change what's lexicalized (emotional conditioning)
- if you grow up in an abusive household, you might not lexicalize 'positive' valence for 'love'
- Though in a sense you might be able to acknowledge that, *for most people* 'love' is a positive feeling
- running a Corpus study would help answer these questions, i think

## 4.3 Playing around: Linguistic tests for content types

Is the affective information targetable by linguistic tests for at-issueness, presupposition

1. Explicit Conversational Rejection
   A: Mary loves me.
   B: No! It's a bad thing!
   $\rightarrow$ Good to my ears, and what you seem to be negating is the Speaker A's (positive) attitude to the proposition.
   but even these tests are defeasable (Simons, Bever, Tonhauser, Roberts collab?)
2. Cancelability
   Mary loves me, and it's a bad thing. $\rightarrow$ Not a contradiction
3. Explicit Propositional Negation
   It's not the case that Mary loves me. $\rightarrow$ does not negate the positive affect
4. Antecedent of the conditional
   in AoC, does S still believe 'love' is positive?
   If Mary loves me, then I'll be upset.
   If Mary loves me, then I'll be happy.
   $\rightarrow$ Both are perfectly felicious. and neither entails that S beleievs that loving is a good thing.

   If the king of france is bald, then there will be a rebellion.
   $\Rightarrow$ S believes there is present king of france.

   If my sister arrives late, I will be upset.
   $\Rightarrow$ I have a sister.

   BUT! Tense matters:
   If my sister arrived late, then I would be upset. But since I don't have a sister, then it doesn't matter (i.e., the vacuously true case where both A and C are false)

5. Bound variable
   Every boy loves his mother
   = where for some of them, loving is good and for some of them loving is bad?
   Context: Sam loves his mother and its positive feeling, Dean loves his mother and it's a negative thing (ew, is it a different sense of love?)

6. Speaker Commitment
   A: Mary love me.
   Did A say that Mary loving A is a good thing?
   Intuition: No.
   $\rightarrow$ The speaker is not committed to the affective content, therefore it is not part of the asserted content.