Psycholinguistics (LING/PSYC 27010)

Autumn, 2016 // University of Chicago

HOMEWORK #4 (due monday, nov14 at 9pm by email)

We need experimental stimuli for the lab session next week! In the first part of this homework, you will build a *minimum of five* items that satisfy the constraints imposed by the design (see below). In the second part, you will identify *at least two* properties of words that could introduce confounds or problems for our interpretation of the results.

Part 1 — stimulus creation

Instructions

Propose five experimental items. For the purposes of this lab, an item consists of a target, a related prime-word, and an unrelated prime-word. Read the details on the next page carefully before starting. Pay attention to the considerations in part 2 while constructing items, as well.

Details for Part 1

We started the experiment-building process on Tuesday by deciding on a design that seems to address our research question pretty well. There will be some tweaks to the design (to be discussed on Tuesday), but the basic question we discussed is still what we're trying to get at.

In class, we considered the possibility of comparing lexical decision times for a word when it is preceded by an unrelated prime-word, versus when it is preceded by a prime-word that is semantically related, but only if both prime and target are understood in an "emerging" (new-ish) internet- or technology-related sense. We hypothesized that a person's age/generation might modulate the priming effect induced by a semantically related prime-word (when compared to an unrelated prime-word).

In order to execute this kind of an experiment, we will need a bunch of target words — to collect decision times on — and a bunch of pairs of prime-words — to manipulate semantic relatedness. So for current purposes, an item is a set of three words `TARGET`, `related`, and `unrelated`, such that:

1. `TARGET` and `related` have a semantic relation that did not exist prior to "the internet;"[1]
2. `TARGET` and `unrelated` are not semantically related in any (reasonable) sense; and
3. `related` and `unrelated` are "comparable" words — roughly, they have a similar number of characters, similar frequencies, perhaps aren't too different morphologically, etc.

The example we discussed in class is listed in the table below, with item id `TL_1` — my initials and a number.

item id	target	related	unrelated
TL_1	TROLL	comment	cement

Following the format of this example, **create five items that satisfy the three constraints listed in points (1), (2), and (3) above, including item id's** (I'll explain why in class). **For at least one item, use words contained in the spreadsheet attached to the email, and indicate the frequency and rank of all three words.** You can find and use an alternative word-frequency database if you track one down yourself and share it with the class.

(part 2 on next page ⟶)

[1] This could be made much more precise, but let's just go with it for the sake of expediency — if the goal was to publish this, we'd have to do better.

Part 2 — fine-tuning the design

If we are not careful about how we choose our stimuli, we introduce the risk of getting patterns in the data that are not directly related to the actual experimental manipulations in our design. For example, suppose we observe that lexical decision times for `TROLL` are shorter (i.e. faster) when `TROLL` is preceded immediately by a 200ms exposure to `comment`, compared to when it is preceded by 200ms of `commutativization`. Does it make sense to attribute this disparity to the semantic relatedness of `TROLL` and `comment`? Probably not, and for many reasons. For one, `commutativization` is far less frequent than `comment`, and contains many more characters. Since frequency and word-length affect recognition speed, we would not really be comparing the relative effects of exposure to `comment` versus exposure to `commutativization` on decision RT for `TROLL`. In fact, 200ms is probably not enough to recognize `commutativization` at all. So we need to make sure that people have enough time to recognize every word, and (ideally) that primes for a common target have comparable frequencies.

Instructions

Identify two features of words (or pairs/sets of words) that we need to pay attention to when selecting our experimental stimuli. For each feature, explain in a few sentences why it matters, given our design. You can follow the model of the `comment-TROLL` versus `commutativization-TROLL` example above, which illustrated the importance of word-length and word-frequency of primes that are presented before a common target word. **Try to be conscious of these features when you are identifying possible items in part 1.**