

“A great deal” versus “a fair deal”: Does collocation strength determine processing speed in aphasia?

Heilemann, C., Zimmerer, V., Varley, R., & Beeke, S.¹

¹*Department of Language & Cognition, Division of Psychology and Language Sciences, University College London*

Introduction

Aphasic language processing is known to be affected by lexical variables such as age-of-acquisition, frequency and length (e.g., Nickels & Howard, 1995). Investigations based on everyday conversational data demonstrate that speakers with non-fluent aphasia make use of common, high-frequency word combinations (i.e. collocations like “*I suppose*”, “*wait a minute*” or “*I know*”; Beeke, 2003; Heilemann, Varley, Zimmerer, Carragher, & Beeke, 2016). This indicates that frequency or familiarity might exert an influence beyond the single word level.

A number of psycholinguistic studies reveal a processing advantage for common, high frequency phrases (e.g., *at the moment*) over matched, less common phrases (e.g., *at the church*) in neurotypical speakers (Arnon & Snider, 2010; Jiang & Nekrasova, 2007; Tremblay, Derwing, Libben, & Westbury, 2011). However, experimental evidence for such phrase frequency or collocation effects in aphasia is relatively rare (e.g., Lum & Ellis, 1999). The present study investigates whether participants with aphasia (PWA) and neurotypical controls show sensitivity to collocation strength in an on-line processing task. We aim to provide insight into the size of units stored within the mental lexicon and explore aphasic language processing from a usage-based perspective.

Methods

Participants

Data from 45 neurotypical control participants (N = 22 in a younger control group, average age = 20.6 years, SD = 3.1; N = 23 in an older control group, average age = 62.7 years, SD = 10.4) and a group of PWA were collected. All participants reported English as their main language.

Materials and procedure

We employ a word monitoring task (WMT), in which the participant reacts (via button press) to a pre-specified target word as quickly as possible once it is encountered in a sentential context (Tyler, Moss, Patterson, & Hodges, 1997). The WMT is an implicit task that reveals, via reaction times (RTs), whether participants are sensitive to targets that appear in varying sentential contexts. Our stimuli consist of pairs of three-word combinations (trigrams). The final word of each trigram pair serves as

the target (nouns such as *PEOPLE*) and is either part of a stronger collocation (e.g., *all the PEOPLE*) or a weaker collocation (e.g., *all three PEOPLE*). All trigrams are derived from the spoken subcorpus of the British National Corpus (BNC, 2007). We use t-scores, a frequency-related measure of the degree of association between the units of an n-gram (in this case a trigram), in order to determine collocation strength. Higher t-scores indicate greater collocation strength (e.g., $t_{all\ the\ PEOPLE} = 10$, versus $t_{all\ three\ PEOPLE} = 2$). All trigrams are embedded in longer, grammatically well-formed sentences which are presented auditorily.

Our WMT comprises three conditions: first-word manipulations (18 pairs, e.g., **round** *the CORNER* vs. **near** *the CORNER*), second-word manipulations (18 pairs, e.g., *a long TIME* vs. *a nice TIME*) and common versus less common noun pairings (CNPs; 12 pairs, e.g., **cats** *and DOGS* vs. **birds** *and DOGS*). CNPs were originally included as distractor items as they are more salient than the trigrams in first- and second-word manipulations. The dependent variable is normalized RT difference (z-score difference) per trigram pair, where a positive z-score difference indicates facilitation of the stronger as compared to the weaker collocation.

In addition to the WMT, participants in the aphasic and the older control group are assessed with a number of cognitive and language tests including a cloze task devised for the purposes of the current study.

Results

Facilitation of word recognition in stronger collocational constructions was found in all conditions in both younger and older neurotypical control participants. Moreover, there was a positive relationship between collocation strength and degree of facilitation in the younger control group, whereas older adults do not show this sensitivity. Preliminary results of seven PWA indicate a shift from normative patterns in that facilitation was found in second-word manipulations and CNPs, but not in first-word manipulations.

Discussion

The current study investigates recognition of words which are embedded in stronger and weaker collocations, derived from naturalistic speech. Facilitation across conditions in neurotypical controls provides evidence for a processing speed advantage of stronger collocations. Collocation strength, as measured by t-scores, determines the degree of facilitation in younger, but not in older controls. This suggests that with more linguistic experience, a threshold (stronger vs weaker) rather than a degree of collocation strength might be in place. In PWA, preliminary evidence indicates facilitated processing in stronger as compared to weaker collocations when the critical word immediately precedes the target, or when the target is primed by a noun (CNPs). However, facilitation in CNPs

might reflect a semantic priming effect rather than an effect of collocation. Relationships between sensitivity to collocation strength and other cognitive and language variables will be reported. Questions about underlying cognitive mechanisms (e.g., whether collocations are stored as chunks in 'lexical' memory, or whether the first word(s) of a collocation prime the rest of it; whether distance or word class govern strength of facilitation) need to be further explored.

References

- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62(1), 67–82. <http://doi.org/10.1016/j.jml.2009.09.005>
- Beeke, S. (2003). "I suppose" as a resource for the construction of turns at talk in agrammatic aphasia. *Clinical Linguistics & Phonetics*, 17(4–5), 291–298.
- Heilemann, C., Varley, R., Zimmerer, V., Carragher, M., & Beeke, S. (2016). Grammatical structures in agrammatism: a usage-based investigation of multi-word expressions. Science of Aphasia XVII, Contributed Papers. *Stem-, Sprak- en Taalpathologie*, 21, Suppl 1, 7-9.
- Jiang, N., & Nekrasova, T. M. (2007). The Processing of Formulaic Sequences by Second Language Speakers. *The Modern Language Journal*, 91(3), 433–445.
- Lum, C., & Ellis, A. W. (1999). Why do some aphasics show an advantage on some tests of nonpropositional (automatic) speech? *Brain and Language*, 70(1), 95–118. <http://doi.org/10.1006/brln.1999.2147>
- Nickels, L., & Howard, D. (1995). Aphasic Naming: What matters? *Neuropsychologia*, 33(10), 1281–1303.
- The British National Corpus*, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>
- Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing Advantages of Lexical Bundles: Evidence From Self-Paced Reading and Sentence Recall Tasks. *Language Learning*, 61(2), 569–613. <http://doi.org/10.1111/j.1467-9922.2010.00622.x>
- Tyler, L. K., Moss, H. E., Patterson, K., & Hodges, J. (1997). The gradual deterioration of syntax and semantics in a patient with progressive aphasia. *Brain and Language*, 56(3), 426–76. <http://doi.org/10.1006/brln.1997.1857>