Project: Word length effects on eye movements of Finnish speakers reading compounds

ESF Exchange Grant, reference number 3792

Grant obtained via: European Network on Word Structure (NetWordS)

Duration visit: 4 months funded by the Exchange Grant (Jan 15th until May 15th, 2012),

2 further months funded by the University of Turku (June/July 2012)

Applicant: Dr. Bernadet Jager, postdoctoral researcher, University of Aberdeen, UK

Address: School of Psychology, William Guild Building, AB24 3FX, Aberdeen, UK

Host: Dr. Jukka Hyönä, University of Turku, Finland

1. Purpose of the visit

The aim of my visit to the laboratory of Dr. Jukka Hyönä at the University of Turku (Finland) was to set up and conduct two eye tracking experiments into the effect of word length on the processing of compound nouns. Bertram and Hyönä (2003) found that Finnish noun-noun compounds are processed holistically when they are short (7-8 letters, e.g. 'jäärata', '*ice track*') but that longer compounds (e.g. 'joukkuehenki', '*team spirit*') are processed sequentially. The proposed studies would consist of a novel combination of a gardenpath paradigm (Staub, Rayner, Pollatsek, Hyönä, and Majewski, 2007) and a word length manipulation (Bertram & Hyönä, 2003): Finnish participants would read sentences containing compound nouns. These compounds as a whole would always be plausible within the sentences, but plausibility of the first constituents was manipulated. Below are examples from the Staub et al. study (italics added).

(a) The maid caught the *elevator mechanic* who was goofing off. (plausible)

(b) The maid scolded the *elevator mechanic* who was goofing off. (implausible)

In (a), the first constituent 'elevator' forms a plausible target of the verb 'caught', so reading should not be impaired, irrespective of whether the compound was processed holistically or sequentially. In contrast, 'elevator' is implausible in (b), and thus provides an opportunity for gardenpathing. Holistic processing would result in similar reading patterns as found for sentence (a), whereas a processing difficulty would indicate that the compound was processed sequentially. In fact, Staub et al. found the latter pattern, indicating that English compounds are processed sequentially, at least when they consist of two nouns that are spaced apart.

The two planned studies would use the same method as Staub et al. (2007). In Experiment 1, Finnish participants would read Finnish sentences, in Experiment 2 Finnish participants would be presented with English sentences. The studies would firstly investigate earlier findings (Bertram & Hyönä, 2003) that long compounds are processed sequentially but short compounds holistically. Secondly, the two experiments allow for comparing processing of Finnish and English compounds by Finnish speakers. Thirdly, the results of these experiments will also open up opportunities to connect research into Finnish and English compounds. For example, the results can be compared to those by Staub et al. (2007), who found that implausible first constituents led to gardenpathing for English compounds, but who investigated spaced nouns and did not manipulate word length.

2. Work carried out during the visit

The original funding proposal was for six months, of which four months were granted. However, Dr. Hyönä has arranged that the remaining two months are being paid for by the University of Turku, allowing me to execute the entire proposal. Originally, the plan was to use the first four months to conduct the Finnish study (Experiment 1), and the remaining two months for the English study (Experiment 2). However, it was discovered that existing Finnish stimulus sets were not suitable for the plausibility manipulation, so the stimulus set had to be created from scratch. This was a time-consuming process. In addition, it was a collaborative undertaking: I constructed the stimuli in English, then Dr. Hyönä translated them into Finnish (more details will be provided below). Therefore, whenever the Finnish stimuli were with Dr. Hyönä, I was able to work on the English stimuli. Currently, both stimulus sets are ready. The next two months (June/July) will be used to conduct one large combined experiment in which 25 participants will first read the Finnish sentences (Experiment 1), and then the English sentences (Experiment 2). Since the summer holiday has already started, participants will be recruited through flyers, mailing lists, and in person. They will be paid with movie tickets, a method that has previously proved a very attractive incentive to participate. Thus, the plan is still to complete the two studies in the originally projected six months.

3. Main results obtained

3.1. Finnish stimulus set

The Finnish target stimuli consist of 60 noun-noun compounds that are also nouns themselves. Half of the words were short (7 to 9 letters, M = 7.7), the remaining thirty words were long (11 to 16 letters, M = 12.2). Part of these words were previously used by Bertram & Hyönä (2003), who obtained them from the Wordmill database by Laine and Virtanen (1999); the remaining stimuli were obtained by translating English compounds into Finnish via Google Translate and Wikipedia, and then looking up the whole words and constituents in the Wordmill database. The same database was also used for obtaining compounding information, lemma frequency (per million) of the whole words and their constituents.

The design was a 2 (word length: short or long) by 2 (sentence context: congruent or incongruent) within-subjects comparison: each participant will be presented with all 60 words, half of them embedded in a sentence that is congruent with both the word as a whole and with the first constituent, half of them incongruent with the first constituent but congruent with the word as a whole; thus 15 stimuli per condition.

Long and short stimuli were matched for whole-word frequency per million ($M_S = 5$; $M_L = 5$; $p = .940$), first-constituent frequency ($M_S = 230$; $M_L = 177$; $p = .367$), and second-constituent frequency ($M_S = 168$; $M_L = 123$; $p = .313$). Frequency ranges were kept similar between the word length conditions for words as a whole ($range_S$ = 0 to 32; $range_L$ = 0 to 46), first constituents ($range_S$ = 2 to 865; $range_L$ = 5 to 780), and second constituents ($range_S$ = 1 to 660; $range_L$ = 0 to 614). Log-transformed scores of raw frequency counts (+1 to eliminate zero scores which cannot be log-transformed) did not differ significantly between short and long words either, neither for words as a whole ($p = .265$), first constituents ($p = .236$), or second constituents ($p = .099$).

The stimuli were embedded in sentences of 7 to 18 words long ($range_S = 8$ to 18; $range_L = 7$ to 15). A t-test showed that number of words did not differ significantly between sentences for short and long target stimuli ($M_S = 10.73$; $M_L = 10.97$; $p = .668$).

For each word, two slightly different sentences were constructed by varying the adjectives preceding the target compound nouns (other than that, the congruent and incongruent sentences were identical). Creation of the sentences was a collaborative process between Dr. Hyönä, a native speaker of Finnish, and myself. Since I don't speak Finnish at all, I translated the target words into English by means of Google Translate, Wikipedia, and several internet dictionaries (Dr. Hyönä checked whether the translation was accurate). I then created the congruent and incongruent sentences, which were translated back into Finnish by Dr. Hyönä. Subsequently, I found more stimuli by doing the reverse: finding Finnish equivalents of English compounds (once more using Google Translate and Wikipedia). Again, the sentences were created in English, then translated into Finnish by Dr. Hyönä.

The first constituents always had higher co-occurrence frequencies with the congruent adjectives than with the incongruent adjectives, as measured by performing Google searches for the combined term ("adjective noun"). For the first constituents, the incongruent word pairs did not occur at all in a large number of cases (18), whereas all congruent word pairs were found. Due to this skewed distribution of what amounts to missing values, no meaningful t-test could be performed for these scores, but the difference was substantial: 19,321 to 146 for the short words and 21,591 to 119 for the long words.

In contrast, the difference was much smaller between the co-occurrence scores of these same adjectives with the whole words (30 and 162 for the short words, $p = .138$; 10 and 142 for the long words, $p = .105$) and the second constituents (1480 and 2416 for the short words, $p = .428$; 1098 and 1364 for the long words, $p = .708$). (For whole words and second constituents, many word pairs were not found either, but this happened similarly often for the congruent and incongruent pairs (whole word = 31 and 28 times; second constituent = 10 and 7 times), thus allowing for statistical comparison.) As a final precaution, the relatively small differences between the congruent and incongruent word pairs went into the opposite direction to that of the first-constituent congruency manipulation: both whole words and second constituents were found slightly more often with the incongruent adjectives than with the congruent adjectives.

The congruent and incongruent adjectives were matched in length, for the short words ($M_{CONGR} = 7.5$; $M_{INCONGR} = 7.8$; $p = .520$) and the long words ($M_{CONGR} = 8.3$; $M_{INCONGR} = 8.7$; $p = .498$). They were also matched for frequency per million, again for the short words ($M_{CONGR} = 140$; $M_{INCONGR} = 83$; $p = .204$) and the long words ($M_{CONGR} = 131$; $M_{INCONGR} = 62$; $p = .089$). Log-transformed frequency scores were significantly higher for congruent adjectives than for incongruent adjectives. However, this was true for both the short (p = .041) and long words (p = .045), so any effect of the word length manipulation would not be due to this variation. In addition, the congruent adjectives were matched between the two length conditions on number of letters (M_S = 7.5; M_L = 8.3; p = .214), frequency per million (M_S = 140; M_L = 131; p = .861), and log-transformed frequency (p = .636). The incongruent adjectives were also matched between the two length conditions on number of letters (M_S = 7.8; M_L = 8.7; p = .107), frequency per million (M_S = 83; M_L = 62; p = .519), and log-transformed frequency (p = .605).

3.2. English stimulus set

The English target stimuli consist of 60 noun-noun compounds that are also nouns themselves. Half of the words were short (7 to 9 letters, M = 8.2), the remaining thirty words were long (11 to 13 letters, M = 11.3). A large number of compound nouns were collected from Gagné & Spalding (2009), Janssen, Pajtas, & Caramazza (2011), Juhasz, Pollatsek, Hyönä, Drieghe, & Rayner (2009), Juhasz, Starr, Inhoff, & Placke (2003), and from the *English Lexicon Project* database (ELP; Balota et al., 2007; http://elexicon.wustl.edu). Words were selected if both the latter and the *Wordsmyth Dictionary and Thesaurus* (WDT; Parks, Ray, & Bland, 1998; http://www.wordsmyth.net) provided an unspaced entry for them (e.g. ‘armchair’ rather than ‘arm chair’), and if the *ELP* database indicated that both the whole words and their constituents were used as nouns. This latter database was also used for obtaining compounding information as well as lemma frequency (including log-transformed frequencies) of the whole words and their constituents.

The design was a 2 (word length: short or long) by 2 (sentence context: congruent or incongruent) within-subjects comparison: each participant will be presented with all 60 words, half of them embedded in a sentence that is congruent with both the word as a whole and with the first constituent, half of them incongruent with the first constituent but congruent with the word as a whole; thus 15 stimuli per condition.

Long and short stimuli were matched for HAL frequency as presented in the *English Lexicon Project* database (ELP; Balota et al., 2007; http://elexicon.wustl.edu). The conditions were matched for whole-word frequency (M_S = 940; M_L = 920; p = .974), first-constituent frequency (M_S = 37,103; M_L = 31,796; p = .614), and second-constituent frequency (M_S = 43,103; M_L = 44,499; p = .913). Frequency ranges were kept similar between the word length conditions for words as a whole ($range_S$ = 9 to 11,306; $range_L$ = 6 to 13,914), first constituents ($range_S$ = 515 to 129,944; $range_L$ = 48 to 156,876), and second constituents ($range_S$ = 1297 to 190,905; $range_L$ = 1214 to 195,199). Log-transformed scores (as presented in the *ELP* database, which uses the LN transformation) did not differ significantly between short and long words either, neither for words as a whole (p = .454), first constituents (p = .999), or second constituents (p = .508).

The stimuli were embedded in sentences of 10 to 20 words long ($range_S$ = 11 to 19; $range_L$ = 10 to 20). A t-test showed that number of words did not differ significantly between sentences for short and long target stimuli (M_S = 14.87; M_L = 14.73; p = .821).

For each word, two slightly different sentences were constructed by varying the adjectives preceding the target compound nouns (other than that, the congruent and incongruent sentences were identical). Creation of the English sentences was mostly done while Dr. Hyönä was working on the Finnish stimuli. Collection of compound nouns had already been done when trying to find additional Finnish target words, since these were obtained by translating English ones.

The congruent adjectives always had higher co-occurrence frequencies with the first constituents than the incongruent adjectives, as measured by performing google searches for the combined term ("adjective noun"). Again, the difference was substantial and similar in scope for the short and long conditions: 571,940 against 58,248 for the short words and 782,827 against 50,672 for the long words. English is much more commonly used on the internet than Finnish, thus both the low and high congruency scores are much higher for the English stimuli than for the Finnish ones; however, the difference between the two conditions is still highly significant for both the short words (p = .010) and the long words (p = .002).

In contrast, the difference was much smaller between the co-occurrence scores of these same adjectives with the whole words (36,438 and 36,806 for the short words, p = .982; 19,767 and 25,321 for the long words, p = .702) and the second constituents (287,806 and 555,821 for the short words, p = .137; 450,956 and 594787 for the long words, p = .572). Again, the differences between the congruent and incongruent word pairs went into the opposite direction to that of the first-constituent congruency manipulation: both whole words and second constituents were on average found more often with the incongruent adjectives than with the congruent adjectives.

The congruent and incongruent adjectives were matched in length, for the short words (M_{CONGR} = 6,3; $M_{INCONGR}$ = 6,6; p = . 504) and the long words (M_{CONGR} = 6.7; $M_{INCONGR}$ = 6.2; p = . 296). They were also matched for frequency per million, again for the short words (M_{CONGR} = 23,187; $M_{INCONGR}$ = 25,709; p = . 837) and the long words (M_{CONGR} = 22,729; $M_{INCONGR}$ = 14,759; p = . 339). Log-transformed frequency scores did not differ significantly between congruent and incongruent adjectives, neither for the short words (p = .074), nor for the long words (p = .115). In addition, the congruent adjectives were matched between the two length conditions on number of letters (M_S = 6.3; M_L = 6.7; p = .432), frequency per million (M_S = 23,187; M_L = 22,729; p = .960), and log-transformed frequency (p = .952). The incongruent adjectives were also matched between the two length conditions on number of letters (M_S = 6.6; M_L = 6.2; p = .333), frequency per million (M_S = 25,709; M_L = 14,759; p = .339), and log-transformed frequency (p = .723).

Future Collaborations with host institution

As previously indicated, my stay has already been extended for two further months. Furthermore, there are plans for future collaborations. Firstly, we are looking into possibilities for a second visit to Turku in 2013. Depending on the results of the current experiments, we would either employ the same stimuli for a modified study, or we could use the extensive collection of collected compound nouns to develop a new set. Secondly, we wish to set up a collaborative project with a research group in an English-speaking country, so that the English stimuli can be tested with native speakers.

Projected publications/articles resulting or to result from grant

The two planned experiments will probably be written up as one combined paper which will discuss the findings in relation to earlier studies in Finnish (e.g. Bertram & Hyönä, 2003) and English (e.g. Staub et al. 2007), and which will also touch upon bilingualism. Furthermore, as discussed above, the stimulus sets may be employed in further studies (in which case the current funding will be acknowledged in resulting publications).

References

Balota, D. A., et al. (2007). The English Lexicon Project, *Behavior Research Methods,* 39, 445-459. Retrieved from: http://elexicon.wustl.edu.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language,* 48, 615-634.

Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language*, 60, 20-35.

Janssen, N., Pajtas, P. E., & Caramazza, A. (2011). A set of 150 pictures with morphologically complex English compound names: Norms for name agreement, familiarity, image agreement, and visual complexity. *Behavior Research Methods,* 43, 478-490.

Juhasz, B. J., Pollatsek, A., Hyönä, J., Drieghe, D., & Rayner, K. (2009). Parafoveal processing within and between words. *The Quarterly Journal of Experimental Psychology,* 62, 1356-1376.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions, and eye fixations. *British Journal of Psychology,* 94, 223-244.

Laine, M., & Virtanen, P. (1999). *Wordmill Lexical Search Program.* Center for Cognitive Neuroscience, University of Turku, Finland.

Parks, R., Ray, J. & Bland, S. (1998). *Wordsmyth English dictionary-thesaurus.* University of Chicago. Retrieved from: http://www.wordsmyth.net.

Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun-noun compounds. *Journal of Experimental Psychology: Learning Memory and Cognition,* 33, 1162-1169.