Readers target words where they expect to find maximal information based on the structure of the lexicon

Jon W. Carr, Monica Fantini, Lorena Perrotti, & Davide Crepaldi

International School for Advanced Studies, Trieste, Italy

Read more: doi.org/k262

SISSA   L2R LANGUAGE LEARNING READING

Do readers look at words based on where they expect to find the most distinctive information?

We formulated a Bayesian cognitive model of visual word recognition to explore how well readers can identify words in different positions. Given a target word t fixated in position j, the reader forms a noisy percept p and then combines this perceptual evidence with prior information about word frequency to form a posterior over the lexicon from which it can ultimately infer a word w^*.
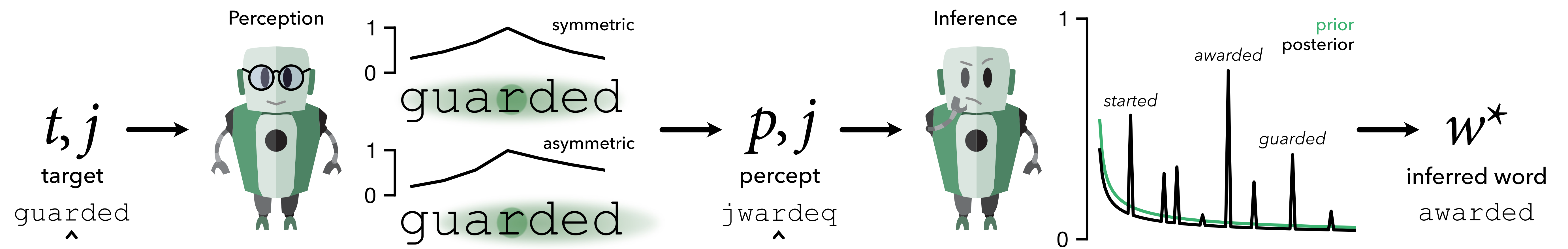

We used the model to look at how information is distributed in natural languages and how this interacts with asymmetries in the visual span...

Accuracy is highest when fixating centrally. However, for left-heavy words such as *guarded*, a left-of-center fixation is preferable to a right-of-center fixation (and vice versa for right-heavy words, such as *concern*). This is further modulated by the shape of the visual span.

Applying the model to data from several languages, we find that there is cross-linguistic variation in information spread.

Target word t: *guarded* (left-heavy word)

Initial fixation, $j=1$	Central fixation, $j=4$	Final fixation, $j=7$
57.8% **guarded**	64.7% **guarded**	23.3% **guarded**
4.8% grandma	7.4% started	14.1% started
3.7% guessed	3.3% quarter	11.4% married
3.0% getting	2.9% married	11.4% decided
2.7% glasses	1.7% learned	2.9% learned
1.7% gunshot	1.5% awarded	1.9% hundred
1.6% granted	0.9% charles	1.8% wounded
1.3% grabbed	0.9% boarded	1.6% changed
1.0% quarter	0.8% charged	1.3% crowded
0.9% goodbye	0.8% grandma	1.3% husband

Symmetric visual span

Initial fixation, $j=1$	Central fixation, $j=4$	Final fixation, $j=7$
79.5% **guarded**	65.2% **guarded**	16.5% decided
2.0% guessed	7.0% started	15.4% married
1.4% grandma	3.1% married	13.0% started
1.2% quarter	2.2% awarded	8.8% **guarded**
1.1% granted	2.1% boarded	3.0% learned
1.0% glasses	1.9% learned	2.5% wounded
0.9% started	1.8% quarter	2.5% changed
0.8% grabbed	1.1% decided	2.4% hundred
0.4% gunshot	1.0% charged	2.1% husband
0.4% learned	0.8% charles	1.5% worried

Asymmetric visual span (right-visual-field advantage)

Target word t: *concern* (right-heavy word)

Initial fixation, $j=1$	Central fixation, $j=4$	Final fixation, $j=7$
53.2% **concern**	80.7% **concern**	72.8% **concern**
6.2% control	1.6% concert	2.2% between
3.3% country	1.0% conceal	1.4% chicken
2.5% college	0.9% concede	1.4% pattern
1.6% contact	0.7% chicken	1.2% popcorn
1.5% concert	0.7% concept	1.0% goddamn
1.3% confess	0.7% dancers	1.0% captain
1.2% concept	0.6% sincere	0.9% western
1.2% chicken	0.6% vincent	0.8% shouldn
1.1% contest	0.5% dancing	0.6% lincoln

Initial fixation, $j=1$	Central fixation, $j=4$	Final fixation, $j=7$
77.5% **concern**	85.4% **concern**	58.2% **concern**
1.8% control	1.1% concert	5.3% between
1.5% concert	0.8% sincere	2.5% pattern
1.2% country	0.7% dancers	2.2% chicken
1.1% college	0.6% chicken	2.1% captain
1.1% conceal	0.5% popcorn	2.1% shouldn
0.8% concept	0.5% conceal	1.7% goddamn
0.7% concept	0.4% lincoln	1.7% western
0.7% condemn	0.3% process	1.6% popcorn
0.6% chicken	0.3% vincent	1.0% kitchen
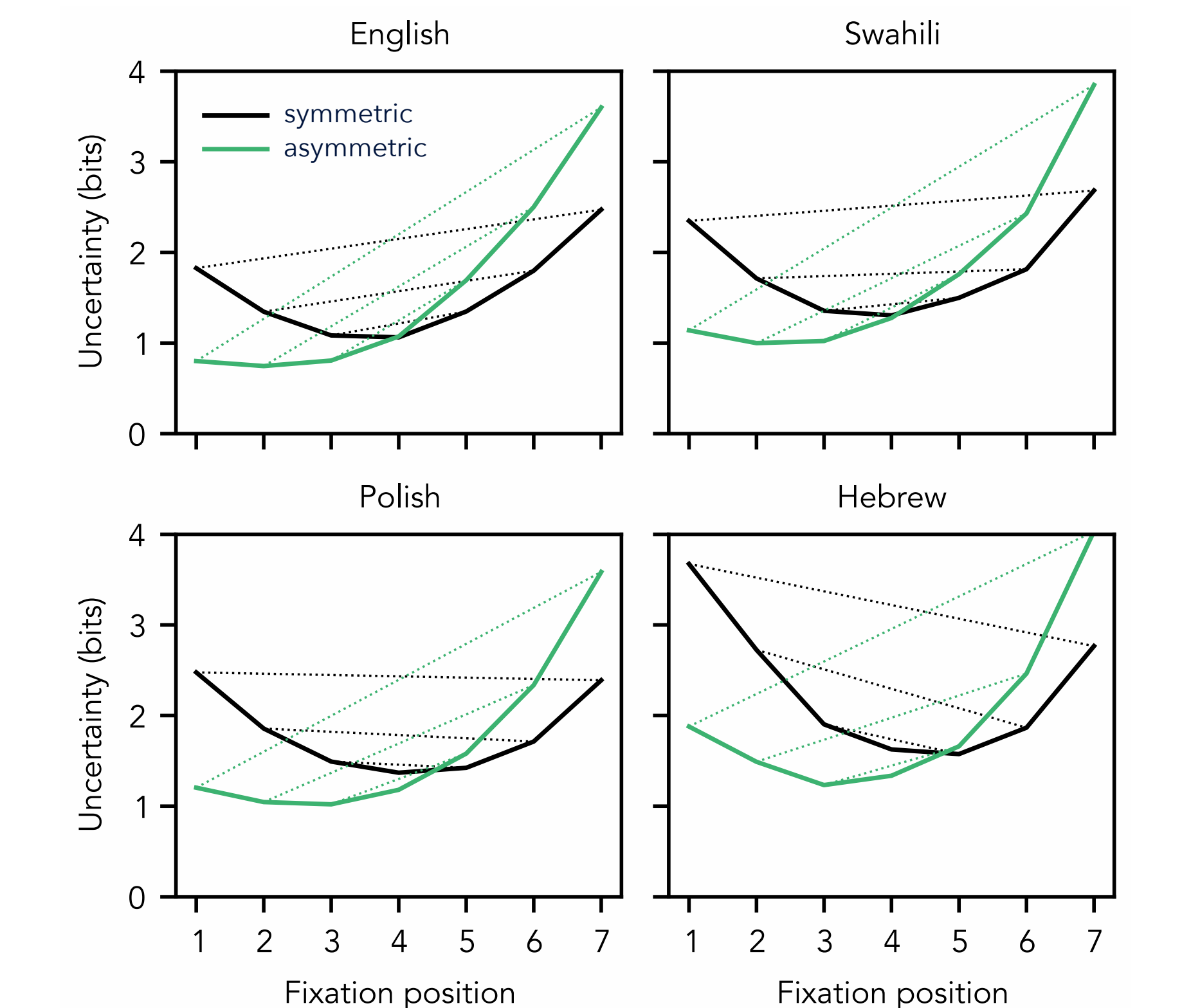
Readers of different languages should target words differently. We tested this prediction using artificial languages in which we can carefully control information spread...

Experiment 1 replicates the classic optimal viewing position effect with artificial lexicons: People more accurately identify words when fixating in more informative positions, modulo perception.

Experiment 2 uses eye tracking to show that people actively target words in positions that minimize uncertainty, taking the perceptual and informational asymmetries into account.

Experiment 3 shows that the effect continues to hold under a more realistic lexicon that contains a mixture of both left- and right-heavy words, albeit with a left- or right-dominance.