The influence of the scene on linguistic expectations: Evidence from cross-modal priming in visual worlds

Andrea Weber & Matthew W. Crocker, Saarland University

aweber@coli.uni-sb.de

INTRODUCTION

- Numerous studies of utterance mediated gaze in visual scenes have demonstrated that sentence processing is not only incremental but also eager: During processing, listeners form expectations about upcoming arguments and make anticipatory eye movements to relevant displayed objects.
- In particular, selectional information from verbs has been shown to guide visual attention to appropriate objects; upon hearing "the boy will eat", listeners start looking at edible objects even before they are mentioned [1, 2].
- While these studies provide valuable insights into semantic processing, it is not clear whether anticipatory eye movements indeed reflect the purely linguistic activation of likely arguments or whether these anticipatory processes are influenced by the circumscribed visual context.
- We present a German cross-modal priming experiment in which we examined listeners sensitivity to selectional restrictions between verbs and their object arguments.

QUESTION

- Do anticipatory eye movements in visual-world studies indeed reflect anticipation of linguistic items in everyday sentence processing or are they merely driven by the presence of depicted objects?

METHOD

- 30 German auditory sentence onsets consisting of a subject NP (the agent) followed by a restrictive verb: "*Die Frau bäckt*" ('The woman bakes')
- Displays with the agent (*Frau*, 'woman'), an object either semantically appropriate as argument for the verb (*Torte*, 'cake') or inappropriate (*Tanne*, 'pine'), and two distractor objects.
- Visual lexical decision items either semantically appropriate (*Pizza*, 'pizza') or inappropriate (*Palme*, 'palm tree') as arguments for the verb. (Always different from depicted arguments.)
- 58 filler trials with varying syntactic structures, restrictive and unrestrictive verbs, cut offs at different locations, or full sentences; with lexical decisions to inappropriate arguments, appropriate arguments, arguments actually mentioned in sentences, or nonwords.
- Ratio of word-to-nonword in lexical decision: 50%.
- Both the appropriate visual argument (pie) and the appropriate lexical decision item (pizza) were highly plausible arguments for sentence onsets (rating study).
- Presentation of items was counterbalanced such that appropriate pictures and lexical decision items also appeared as inappropriate items ("*Die Frau pflanzt*", 'The woman plants').
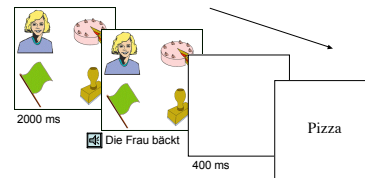

EXPERIMENT 1

- 48 German listeners
- Task: listen to sentences while seeing pictures and respond to subsequent visual lexical decision items
- Auditory: The woman bakes
- Visually: pie or pine
- Lexical decision: pizza or palm tree
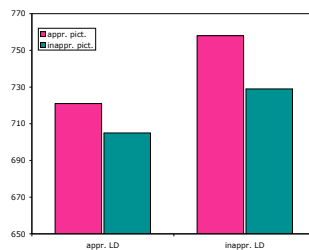
- As expected, reaction times were faster for lexical decision items which were semantically appropriate (*pizza*) than for inappropriate items (*palm tree*; $F_1[1, 47]$ = 8.54, $p < .01$; $F_2[1, 29]$ = 4.24, $p < .05$).
- Surprisingly, however, reaction times were slowed when the display included a picture of an appropriate argument (*pie*) prior to lexical decision compared to when no appropriate argument was depicted (*pine*). This effect was significant by subjects ($F_1[1, 47]$ = 5.03, $p = .03$; $F_2[1, 29]$ = 2.04 $p > .05$).
- The semantic interference from pictures occurred both when the lexical decision item was appropriate and when it was not (no interaction: F_1 & $F_2 < 1$).
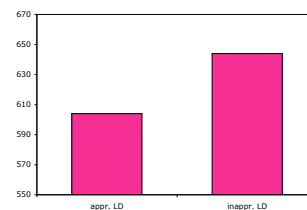
EXPERIMENT 2

- To determine the effect of additional picture primes on lexical decision times, only auditory primes were included in Experiment 2.
- 24 German listeners
- Task: listen to sentences and respond to subsequent visual lexical decision items (no pictorial prime)
- Auditory: The woman bakes
- Lexical decision: pizza or palm tree

- As in Experiment 1, reaction times were fast for lexical decision items which were semantically appropriate (*pizza*) than for inappropriate items (*palm tree*; $F_1[1, 23]$ = 6.95, $p < .02$; $F_2[1, 29]$ = 3.99, $p = .05$).
- The presence of picture primes in Experiment 1, resulted in significantly slower reaction times compared to Experiment 2, both when the picture primes were appropriate or inappropriate.

GENERAL DISCUSSION

- Facilitated lexical decision times for appropriate items in Experiment 1, regardless of the scene, provide evidence for purely linguistic anticipation, confirming the gated completion findings of Altmann (1999).
- In contrast to previous priming results (see e.g. [3]), however, appropriate picture primes slowed down lexical decision times. We suggest, that visually attending the picture of an object based on supporting auditory input, set up expectations for that specific object to follow as verb argument; when the expectations were not met, lexical decision times were slowed.
- To test whether indeed the combination of auditory primes with picture primes was responsible for the semantic interference, we are currently re-running the study with just picture primes. Also the presence of four picture primes rather than a single picture prime might have influenced the result patterns. (Indeed, lexical decision times in Experiment 2 were always faster than in Experiment 1.) We will therefore compare lexical decision times following single picture primes and multiple picture primes.

REFERENCES

[1] Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 72.

[2] Kamide,Y., Altmann,G.T.M., & Haywood,S.(2003).The time course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. *Journal of Memory and Language*, 49.

[3] McQueen, J. & Huettig, F. (2005). Semantic and phonological priming of auditory lexical decision by pictures and printed words. *AMLaP*, Ghent, Belgium.

[4] Altmann, G.T.M. (1999). Thematic role assignment in context. *Journal of Memory and Language*, 41.