

Immediate effects of anticipatory coarticulation in spoken-word recognition

A presentation by George Daniel Mikleu,
based on the same named paper of Salverda,
Kleinschmidt and Tanenhaus (2014)

Coarticulation

- Both within and between words
- 2 Types:
 - Perservative: Utterance production influenced by previous utterance(s)
 - Anticipatory: Utterance production influenced by following utterance(s)
 - Pre-onset anticipatory coarticulation focus of study
 - Thought well-documented, not much empirical evidence in case of concurrent speech processing

Time Lag

- Time lag between speech and eye movements suspected
- Generally 200 ms assumed
- often affects result interpretation of visual-world studies

Altmann (2011)

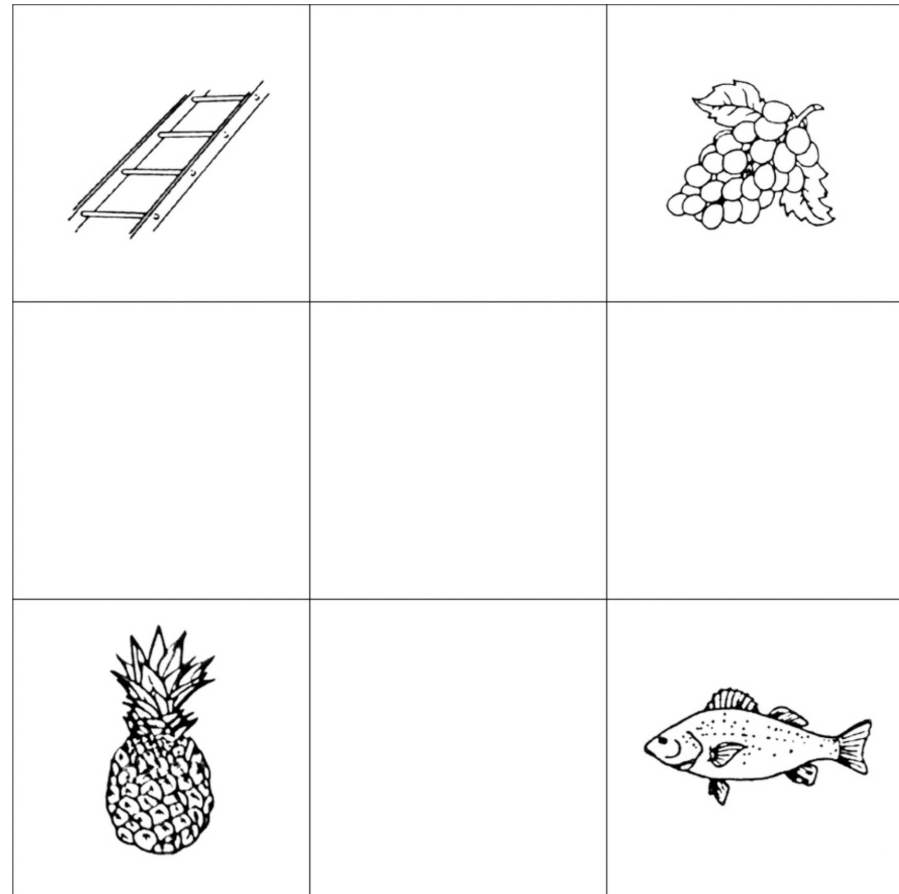
- Proposes time lag to be no more than 100 ms
- Analyzed sentences all began with “the”
 - Strong coarticulatory effects of schwa (/ə/) expected
- Rejects explanations based on coarticulation:
 - Why not similar results in other visual-world studies?
 - Evidence of 200 ms time lag for within-word anticipatory coarticulation (Dahan et al., 2001)

The Experiments

- 2 visual-word experiments
- Experiment 1: Searches minimal time until acoustic information utilized
- Experiment 2: Looks at listener's usage of pre word-onset anticipatory coarticulation

Experiments – Materials

- 18 visual displays, 4 pictures each: 2 distractors, 2 potential targets
- Potential targets:
 - Real target and additional “associated distractor” (switched equally between trials)
- 11 additional displays for practice (5) and fillers (6)



Experiments – Materials

- Different first vowel for each picture name
- Distractors:
 - Onset articulation differed from both potential targets
- Potential targets:
 - One has approximant (/r/, //, or /w/), other voiceless fricative (/f/, /s/, or /ʃ/) in the beginning

Potential targets Approximant- initial	Fricative- initial	Distractor	Distractor
ladder	fish	grapes	pineapple
lamp	shirt	mountain	bus
leaf	foot	bottle	cannon
lemon	fly	pot	garbage can
lion	shoe	comb	pen
lobster	flag	pipe	bed

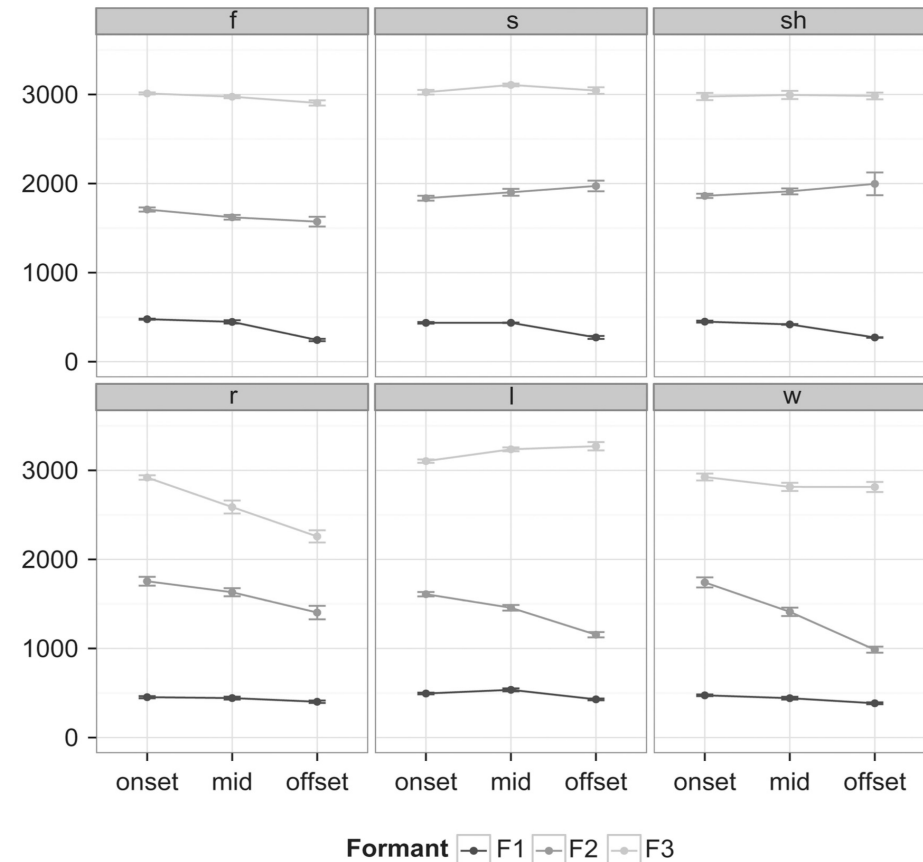
Experiments – Materials

Recorded sentences:

- “the” only relevant determiner
- Experiment 1: determiner and target word separated by prosodic break
 - Example: “The ... ladder is the target”
 - Original determiner replaced by isolated recording
- Experiment 2: determiner and target word continuous, followed by prosodic break
 - Example: “The ladder ... is the target”

Coarticulatory Information of the Determiner

- First 3 formant frequencies at onset, midpoint and offset of schwa vowel
- F1 drop at schwa offset for approximants, even further for fricatives
- F2 consistent rise for /s/ and /ʃ/ (sh); significant drop for approximants
- F3 relatively flat throughout schwa for fricatives, but distinct pattern for each approximant

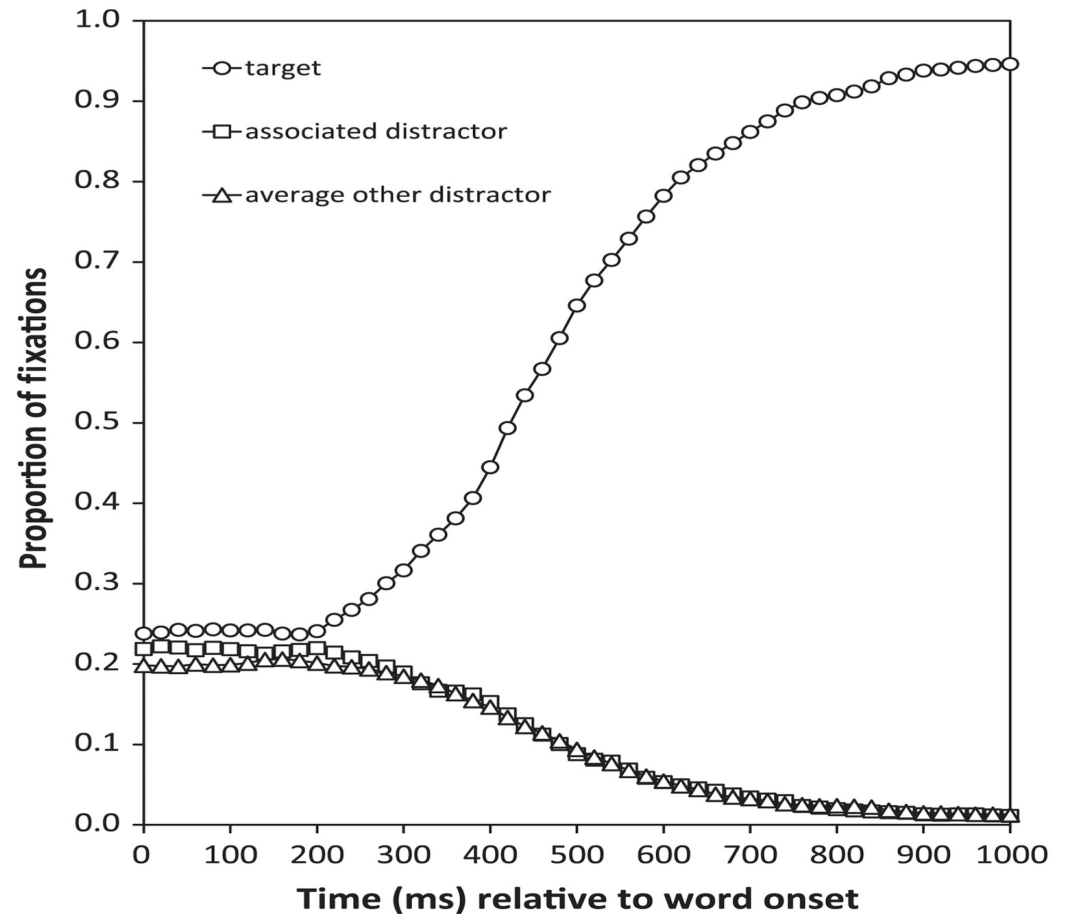


Experiments – Design & Procedure

- Each experiment 60 participants:
 - Experiment 1: 41 women; 18 – 34 years (mean: 20)
 - Experiment 2: 45 women; 18 – 24 years (mean: 19.7)
- 15 random trial orders combined with 4 visual display lists to 60 randomized lists
- Should click pictures indicated by respective sentence
 - Eye-movement recorded
 - Every 5 trials: fixation dot in center

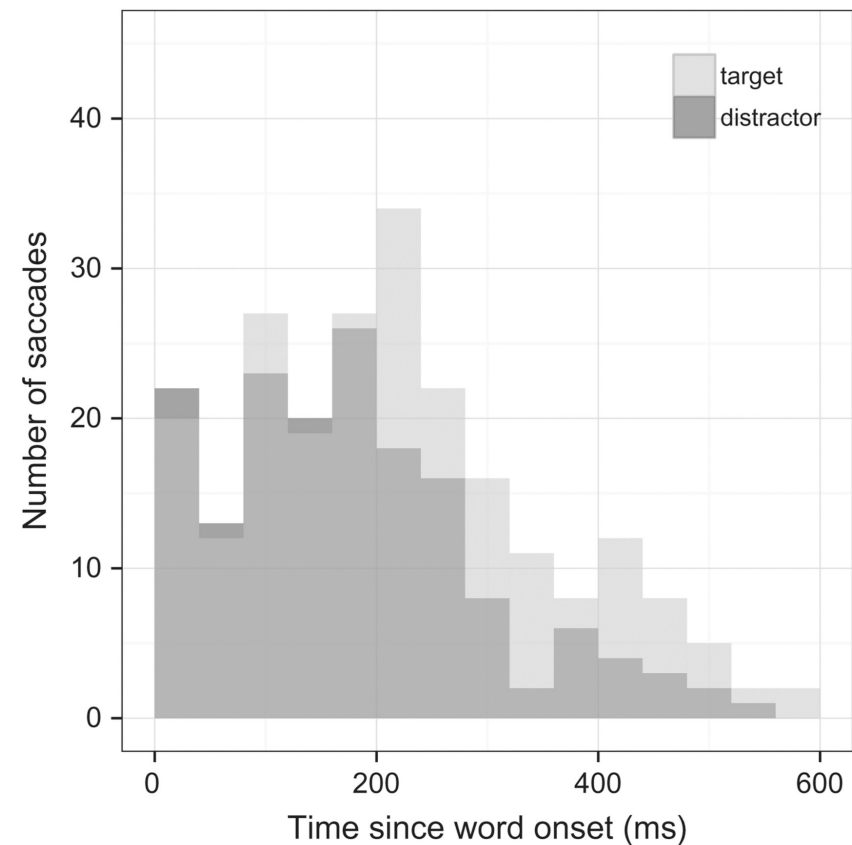
Experiment 1 – Fixation Proportions

- **Fixation Proportions:**
 - Till 200 ms: no significant differences
 - After 200 ms: Fixation proportion for target increased, while other proportions decreased



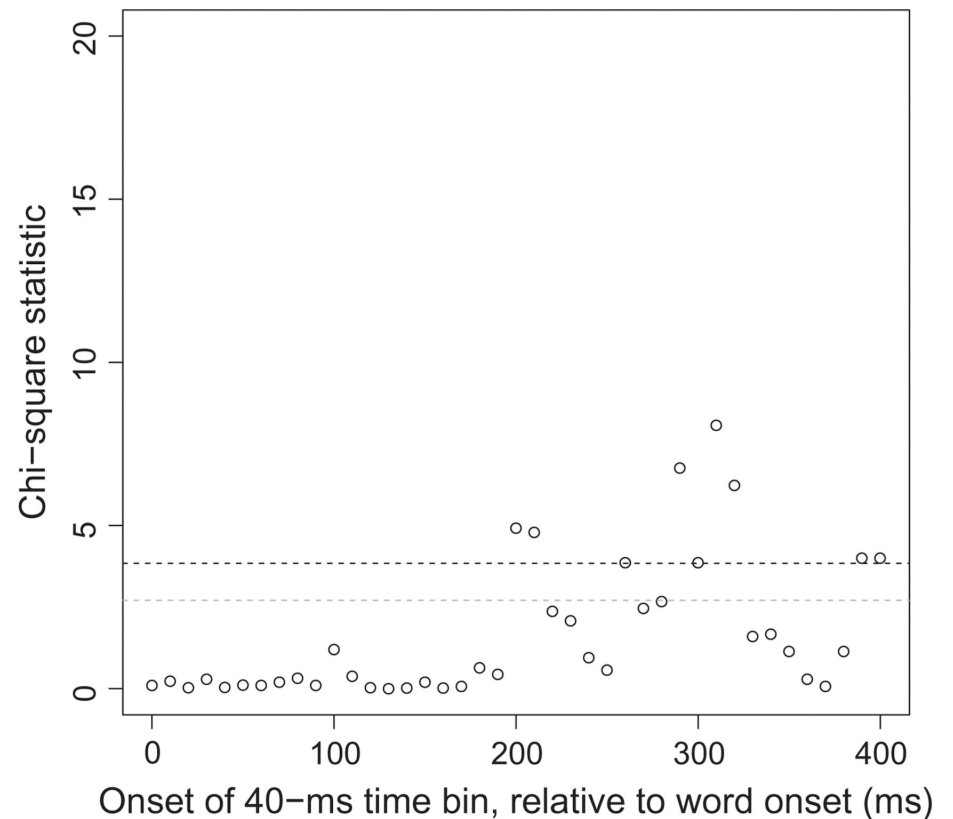
Experiment 1 – Saccade Latencies

- Till 200 ms: no particular tendencies
- After 200 ms: always more trials with saccade to target than to associated distractor



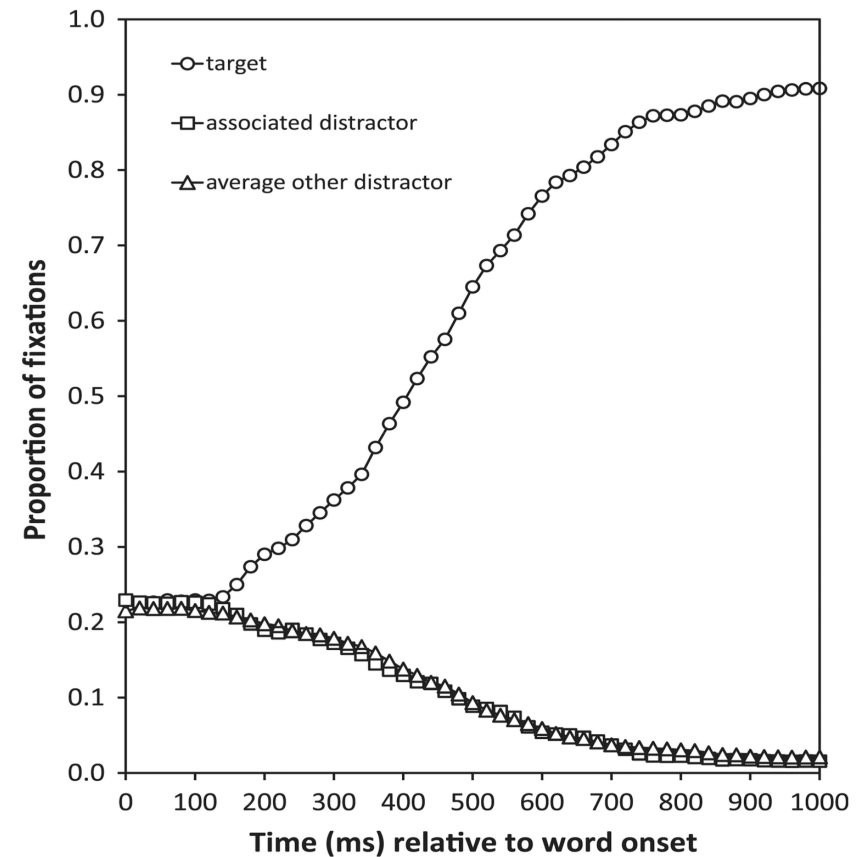
Experiment 1 – Saccade Latencies: Moving-Window Analysis

- Till 200 ms: no significant differences
- Significant effects in 200–240 ms and 210–250 ms bins and around 300 ms and 400 ms after word start
- **Conclusion:** No significant effects of speech signal on eye-movement control before 200 ms after target start



Experiment 2 – Fixation Proportions

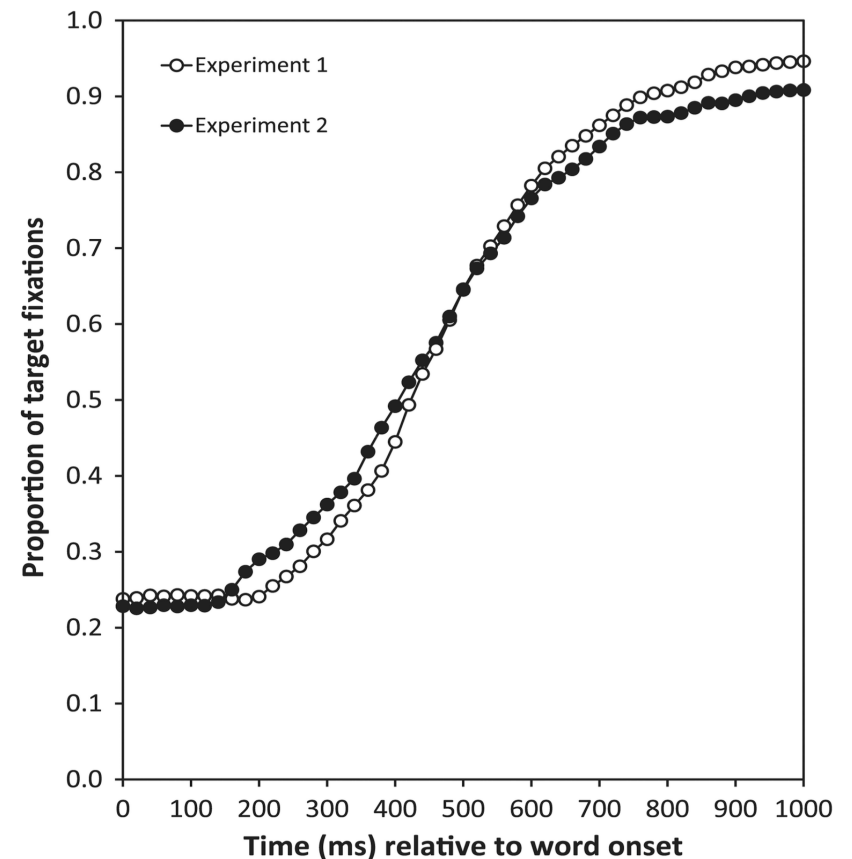
- Till 100 ms: no significant difference
- 100 - 200 ms: Fixation proportion for target increased, while other proportions decreased



Comparisons

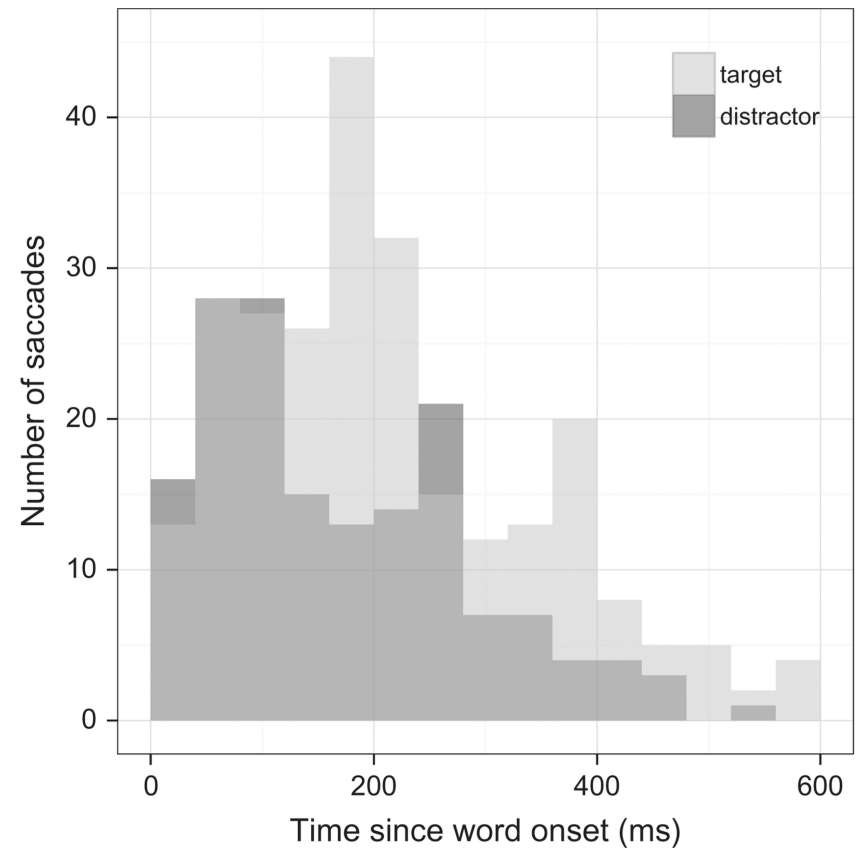
Experiment 1 / Experiment 2

- Target fixations increased earlier in Experiment 2
- Target fixations reached higher level in Experiment 1
 - likely because of greater word length in Experiment 2 (due to prosodic break)



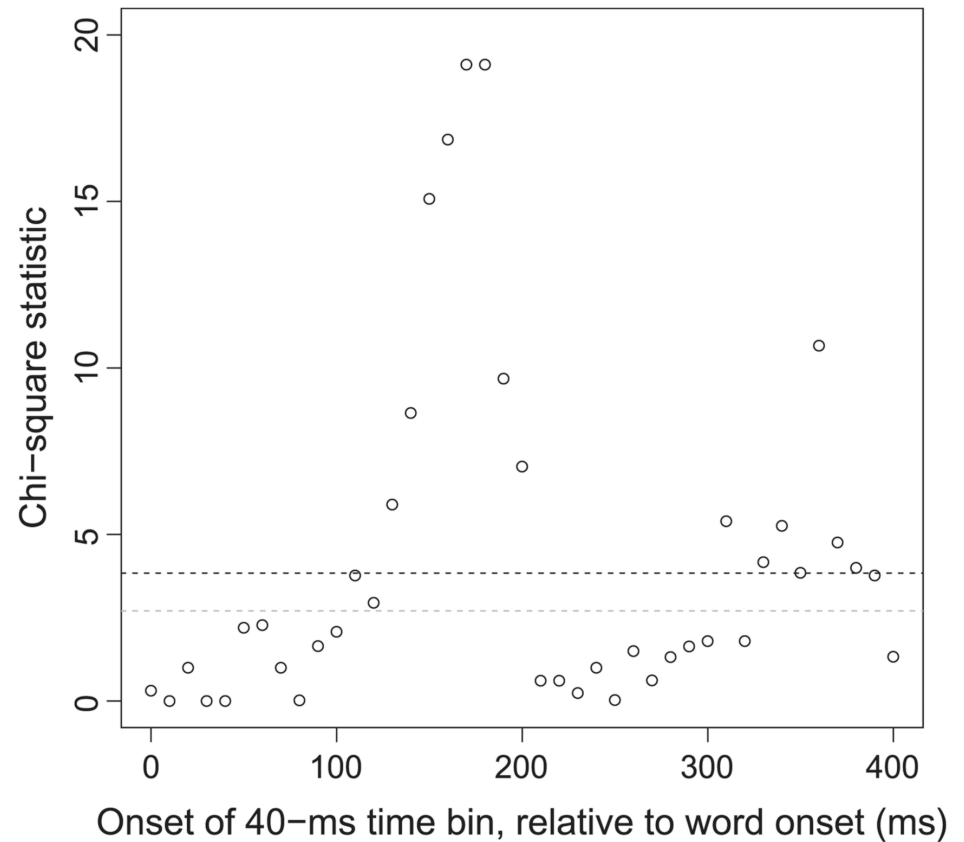
Experiment 2 – Saccade Latencies

- Till 120 ms: no particular tendencies
- After 120 ms: always more trials with saccade to target than to associated distractor (only exception: 240 - 280 ms time bin)



Experiment 2 – Saccade Latencies: Moving-Window Analysis

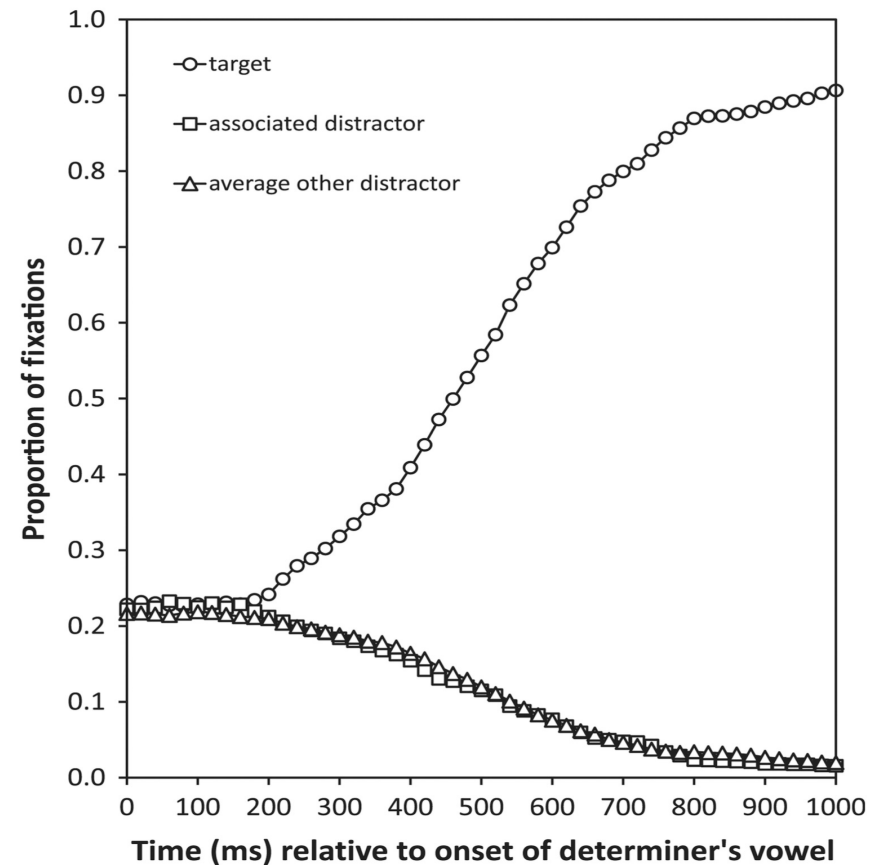
- Significantly more saccades on target for bins starting at 130-240 ms: **about 70 ms earlier in Experiment 2 than in Experiment 1**
- Significant effects in most bins starting at 300-400 ms



Experiment 2 – Fixation Proportions

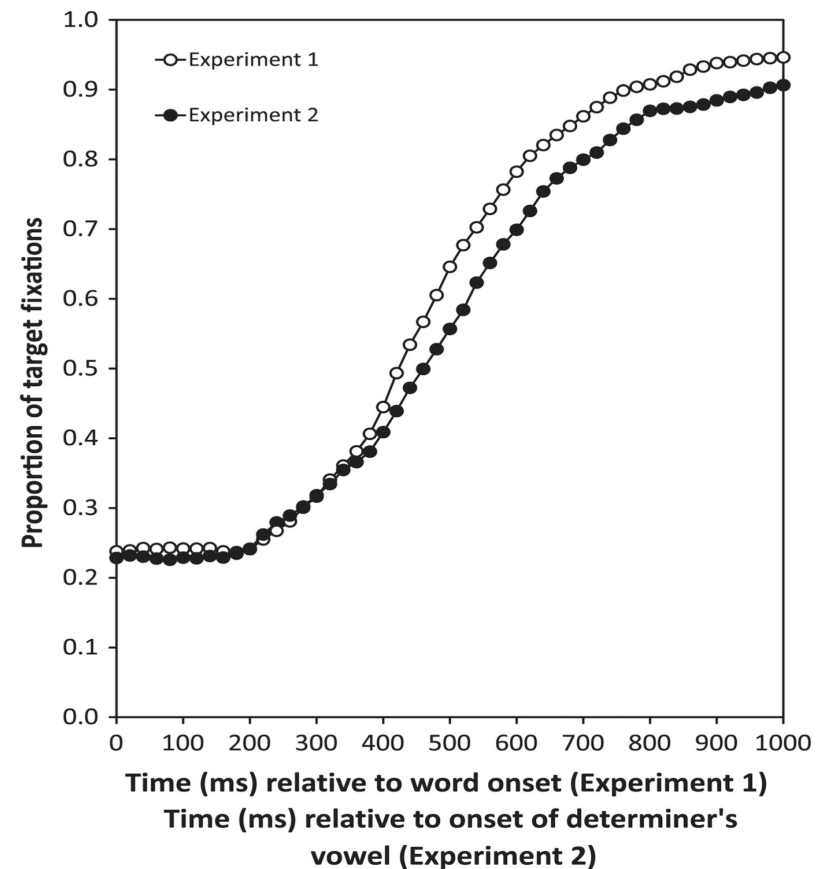
(relative to determiner vowel onset)

- Anticipatory coarticulation suggests fixations to be aligned to start of determiner vowel rather than to noun start
- After reanalysis rise of fixation proportion for target only slightly before 200 ms



Comparisons Experiment 1 / Experiment 2 (relative to determiner vowel onset)

- After reanalysis target fixations start rising at about 200 ms in both
- **Conclusion:** Listeners (immediately) use coarticulatory information to predict beginning of next word
 - Additionally, support for 200 ms time window



Experiment 2 compared to Altmann (2011)

- Experiment 2: No significant effects 80-120 ms after word onset
 - Earliest: 130-170 ms time bin
- Exact effect timings not directly comparable, however:
 - Differences in stimuli, design, procedure, and experimental task between studies
 - Determiners in Altmann's studies longer than here
 - Information earlier available, relative to target word onset

Methodological implications for visual-world experiments

- Results support 200 ms time lag assumption
- Affirm many previous visual-world studies with this assumption
- Altmann's suspected 100 ms minimum lag likely because of anticipatory coarticulation preceding estimated noun onset by about 80–100 ms
- Why not more reports of smaller time lags in visual-world studies?
 - not enough details to observe early effects of anticipatory coarticulation

Implications for models of spoken-word recognition

- Results can generally be explained by either pattern recognition, exemplar or explanation-based models
- Pattern-Recognition Models:
 - Could explain coarticulation by enriching phonetic detail, for example with context sensitive features
- Exemplar Models:
 - Could explain coarticulation by letting exemplars represent multiple word utterances
- Data-Explanation Models:
 - Could explain coarticulation by hypotheses about upcoming word/sound

Implications for models of spoken-word recognition

- Greatest support for explanation-based models:
 - Can additionally explain flexibility of anticipatory coarticulation across speakers and linguistic contexts
 - Other model types often struggle, if meaning of acoustic patterns can change with context

Implications for models of spoken-word recognition

- Additionally, interesting suggestions for communication (English)
- Phonetic properties of function words potential consequence of assisting in:
 - processing neighboring speech sounds
 - recognizing subsequent content words

References:

- Altmann, G. T. (2011). Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. *Acta psychologica*, 137(2), 190-200. <https://doi.org/10.1016/j.actpsy.2010.09.009>
- Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, 16(5–6), 507–534. <https://doi.org/10.1080/01690960143000074>

References:

- Salverda, A. P., Kleinschmidt, D., & Tanenhaus, M. K. (2014). Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of memory and language*, 71(1), 145-163. <https://doi.org/10.1016/j.jml.2013.11.002>