

PROBABILISTIC REDUCTION IN READING ALOUD: A COMPARISON OF YOUNGER AND OLDER ADULTS

Cornelia Moers^{1,2}, Esther Janse^{3,4,1}, & Antje Meyer^{1,4}

¹ Max Planck Institute for Psycholinguistics; ² Max Planck International Research Network on Aging; ³ Centre for Language Studies, Radboud University Nijmegen, ⁴ Donders Institute for Brain, Cognition and Behaviour
Email: cornelia.moers@mpi.nl

ABSTRACT

Frequent and predictable words are generally pronounced with less effort and are therefore acoustically more reduced than less frequent or unpredictable words. Local predictability can be operationalised by Transitional Probability (TP), which indicates how likely a word is to occur given its immediate context. We investigated whether and how probabilistic reduction effects on word durations change with adult age when reading aloud content words embedded in sentences.

The results showed equally large frequency effects on verb and noun durations for both younger ($M_{\text{age}} = 20$ years) and older ($M_{\text{age}} = 68$ years) adults. Backward TP also affected word duration for younger and older adults alike. Forward TP, however, had no significant effect on word duration in either age group.

Our results resemble earlier findings of more robust Backward TP effects compared to Forward TP effects. Furthermore, unlike often reported decline in predictive processing with aging, probabilistic reduction effects remain stable across adulthood.

Keywords: Word durations, Reading aloud, Aging, Frequency effects, Predictability

1. INTRODUCTION

The *Probabilistic Reduction Hypothesis* [11] states that more probable words are acoustically more reduced than less probable words. The probability of a word is influenced by its frequency (a word's prior probability) and its frequency in specific contexts (a word's contextual probability). The latter can be measured by Transitional Probability (TP), which indicates the probability of a word given its right or left neighbouring word. Word frequency and transitional probability have been shown to influence acoustic reduction (involving segment deletion or weakening and shortened word duration) in several corpus studies of conversational speech [e.g., 5, 14, 9].

Apart from their influence on acoustic properties in spontaneous speech production, word frequency

and transitional probability influence reading behaviour. This has been shown in several studies on frequency and predictability effects on eye movements [7, 13]. The more likely a word is to occur (given a specific context), the less time the eyes spend on it and the more easily that word can be processed.

Whereas frequency and transitional probability effects have been replicated in several studies investigating silent reading, little is known about these effects on spoken word durations and eye movements in reading aloud. Reading aloud is a complex and demanding task involving the orchestration of several subcomponents: The visual recognition of words, the planning of speech and the actual articulation of the read words need to be synchronized efficiently [1]. Therefore, articulation of a word while reading aloud may differ from spontaneous speech production in that articulation can be influenced by orthographic properties of the text such as punctuation, and by simultaneous visual processing of words further ahead in the text [10].

Two studies used a reading-aloud task when investigating probabilistic reduction effects: Hanique and Ernestus [9] investigated predictability effects on final /t/-reduction in different speech registers (conversational speech, semi-formal interview speech and formal speech, i.e., reading aloud). They found probabilistic reduction effects for all registers, but reduction effects were larger for more spontaneous speech. Secondly, Baker and Bradlow [2] compared effects of second-mention reduction in plain reading aloud to those in clear reading aloud. Reduction effects occurred in both reading styles, but effects were again larger in the less formal speaking style. Note, that neither of these studies investigated effects of TPs between two words on word durations. Thus, there is only limited evidence that probabilistic reduction effects transfer to reading aloud. Knowledge about these effects would give important insights for advancing theories of probabilistic reduction in speech as well as developing models of reading aloud.

Importantly, thus far, all studies on probabilistic reduction effects have investigated student populations or middle-aged adults. However, probabilistic reduction effects may change in size with advanced adult age. Age has been found to

modulate word frequency effects in several studies of silent reading or speech production. However, results of these studies did not converge: On the one hand, older adults have been reported to show larger word frequency effects than younger adults in silent reading [17]. On the other hand, increased language experience has been shown to lead to decreased frequency effects in picture naming [8]. Studies investigating context effects have also yielded inconsistent age effects: Older adults show either smaller context effects in speech comprehension [6], or make equal use of context to predict upcoming words in silent reading as a younger group [17]. Furthermore, it is unclear whether age differences in frequency and predictability effects found in silent reading, speech comprehension or picture naming transfer to probabilistic reduction effects in read-aloud speech.

In the present study, we investigated the following questions: First, do probabilistic reduction effects as found in conversational speech generalise to reading aloud? Second, how do probabilistic reduction effects develop across adult age?

2. METHODS

To study the effects of word frequency and transitional probability (TP) on reading behaviour in younger and older adults, we designed an eye-tracking experiment that included two reading tasks: silent reading and reading aloud. The speech recordings from the oral reading condition are analysed in the current study.

2.1. Participants

Thirty older and thirty younger adults participated in the experiment (recruited via the participant pool of the Max Planck Institute for Psycholinguistics) and were paid for their participation. Older participants' mean age was 68.37 years ($SD = 4.49$; range: 62 to 78 years), younger participants' mean age was 20.83 years ($SD = 2.70$; range: 18 to 27 years). All participants were native speakers of Dutch, with no self-reported history of language impairments or neurological problems. Furthermore, all participants had normal or corrected-to-normal vision, which was tested prior to the reading experiment to ensure that participants could properly read from the computer screen.

2.2. Materials and procedure

Testing and recording took place in a sound-treated booth at the Max Planck Institute for Psycholinguistics. At the start of a session,

participants received written information about the upcoming reading task and gave written consent. Participants were instructed to read sentences from a computer screen. They saw two example sentences which they had to read out loud in order to familiarise themselves with the task. Sentences were presented on a single line in an 18 point font (Times New Roman) as black letters on a white background, vertically centred, but horizontally left-aligned. After participants had finished reading a sentence, they had to press a mouse button to proceed to the next sentence. An untimed yes/no comprehension question occurred after one quarter of the sentences. Response accuracy for the comprehension questions was generally high (mean accuracy being 96% for the younger and 93% for the older group).

Participants read 240 Dutch sentences in total, divided over the silent and reading aloud conditions, which alternated in four blocks (block 1: 60 silent trials, block 2: 60 aloud trials, block 3: 60 silent trials, block 4: 60 aloud trials; or reverse order). Each sentence included a target noun-verb combination. Transitional probabilities for these combinations, as well as word frequencies for each noun and verb, were drawn from the 44 million words SUBTLEX-NL corpus [12]. Half of the 240 items were infinitives (e.g., “afspraak maken” – *to make an appointment*) and the other half were past participles (e.g., “fouten gemaakt” – *made mistakes*). Forward transitional probability indicates the predictability of the verb from the noun and was computed by taking the total number of occurrences of the noun-verb combination divided by the number of occurrences of the noun. Log-transformed forward transitional probabilities for the 240 noun-verb combinations followed a uniform distribution and thus varied continuously from very high log probability (e.g., $\log FTP = -0.31$ for “afscheid nemen” – *to say farewell*) to very low probability (e.g., $\log FTP = -9.88$ for “tijd kiezen” – *to pick a time*). For each noun-verb combination, we also calculated backward transitional probability, being the predictability of the noun given the verb.

The noun-verb combinations were embedded in neutral, non-predictive sentences that did not contain words that were semantically related to the upcoming target noun or verb. Sentence length ranged from seven to twelve words (including the noun-verb combination). Sentences were constructed such that the target combinations were neither sentence-initial nor sentence-final to avoid initial or final lengthening effects.

2.3. Annotations

Participants' speech recordings were annotated by one of eight experienced transcribers in order to

derive total sentence reading times and word durations (in milliseconds) for each target noun and verb. Annotations were done using the software Praat [15]. Disfluent or misread sentences were excluded from the analysis. A few sample sentences of each transcriber and each participant were evaluated for accuracy by the first author. Annotation accuracy was evaluated as being very high. Speech rate for each sentence was measured by dividing the total sentence reading time by numbers of syllables in the orthographic sentence representation.

3. RESULTS

3.1. Statistical Analysis

To investigate probabilistic reduction effects on word durations, we fitted linear mixed-effects models in R version 3.1.1. [16], using the lmer-function from package lme4 [4]. Each model included word length (in number of characters for either the verb or noun), speech rate (in syllables per second, log-transformed), and verb type (infinitive or past participle) as control variables, as well as word frequency (log frequency per million words of either the verb or noun), transitional probability (either log forward or log backward TP) and age group (older versus younger group) as predictor variables. Word durations were log-transformed. Numerical predictors were centered around their mean to reduce non-essential collinearity [19]. All regression models contained participants and items as random effects, as well as by-subject random slopes for frequency and TP (uncorrelated) and by-item random slopes for age group [3, following a design-driven approach]. P-values were obtained by likelihood-ratio tests (comparing nested models). Since all predictor and control variables were theoretically motivated, no model stripping procedure was applied. As our focus was on possible age-related changes in the effect of predictability after controlling for effects of word frequency, we did not remove collinearity between the frequency and TP variables by residualisation [19, for a discussion]. Instead, we always fitted an initial model that solely investigated the effect of frequency (and its interaction with age group), and only afterwards added either ForwardTP or BackwardTP (to the fixed and to the random part). No changes in frequency effects or frequency by age group interactions occurred in any model due to the addition of the TP variable. The models reported below are based on 6543 trials (out of 7200 trials in total, i.e., 9.1% data loss due to exclusion of disfluent and misread sentences). Older adults took overall more time to read sentences ($M = 2797$ ms, $SD = 548$) than younger adults ($M = 2429$ ms, $SD = 442$), and thus

older adults had slightly slower speech rates ($M = 5.14$ syllables/second, $SD = 0.85$, Range = 2.47-11.33) than younger adults ($M = 5.86$ syllables/second, $SD = 0.91$, Range = 2.47-11.33). Consequently, older adults had longer verb durations (OA: $M = 493$ ms, $SD = 142$, Range = 149-1198) than younger adults ($M = 414$ ms, $SD = 119$, Range = 90-1037), as well as longer noun durations (OA: $M = 398$ ms, $SD = 117$, Range = 139-954; YA: $M = 349$ ms, $SD = 97$, Range = 110-743).

3.2. Verb duration

A summary of the best-fitting linear mixed-effects model for the verb duration analysis is shown in Table 1. The model showed significant simple effects of word length, speech rate, verb type, and verb frequency, as well as a significant interaction of ForwardTP by age group. As expected, orthographically longer words had longer spoken durations, higher global speech rate led to shorter word durations, and verbs that occur often were acoustically more reduced than verbs with low frequency of occurrence. Past participle verbs had longer durations than verbs of the infinitive type. Furthermore, the older adults had longer word durations than the younger adults. ForwardTP showed no influence on verb duration for the younger adults (i.e., the group mapped on the intercept). Older adults reduced verbs more when they were more predictable from the noun than younger adults, as indicated by the age group by ForwardTP interaction. However, this interaction is difficult to interpret, as rerunning the same model with the older group mapped on the intercept did not show an effect of ForwardTP either. Hence, the effect of ForwardTP is overall not strong enough to significantly influence word duration.

Table 1: Estimates, standard errors and t values for mixed-effects model fitted for the verb duration data; significance ($p < 0.05$) indicated by asterix (*).

Variable	β	SE	t
WordLength	0.09223	0.00540	17.1*
LogSpeechRate	-0.55765	0.01640	-34.0*
VerbForm (PP)	0.07588	0.01598	4.7*
LogFreqVerb	-0.03622	0.00955	-3.8*
LogForwardTP	0.00281	0.00404	0.7
Group (Older)	0.09519	0.01846	5.2*
LogFreqVerb by Group	0.00295	0.00418	0.7
LogForwardTP by Group	-0.00393	0.00187	-2.1*

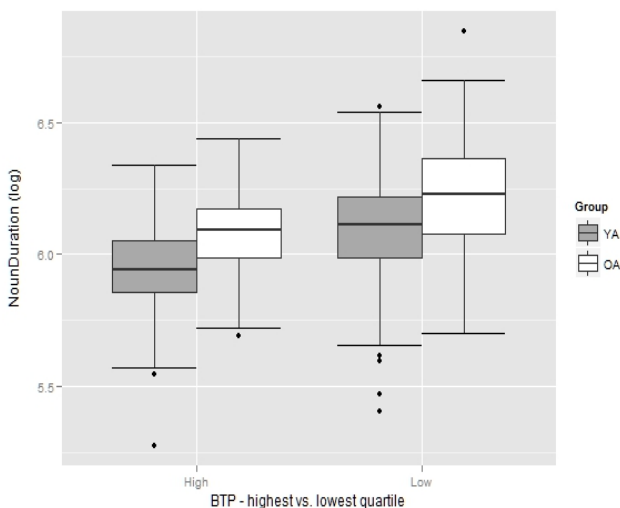
3.3. Noun duration

A summary of the best-fitting linear mixed-effects model for the analysis of noun durations is shown in Table 2. The model showed significant simple effects of word length, speech rate, noun frequency, and BackwardTP, but no significant interactions between either noun frequency and age group or Backward TP and age group. As expected, longer words had longer durations, higher speech rate led to shorter word durations, and frequent nouns were acoustically more reduced than infrequent nouns. Moreover, the predictability of a noun from the verb influenced noun durations such that more predictable nouns were more reduced than less predictable nouns. This effect occurs for younger and older adults alike, as indicated by the non-significant age group by BackwardTP interaction. Figure 1 illustrates age group and TP effects.

Table 2: Estimates, standard errors and t values for mixed-effects model fitted for the noun duration data; significance ($p < 0.05$) indicated by asterix (*).

Variable	β	SE	t
WordLength	0.10577	0.00523	20.2*
LogSpeechRate	-0.45300	0.01682	-26.9*
VerbForm (PP)	-0.01560	0.01792	-0.9
LogFreqNoun	-0.02515	0.01006	-2.5*
LogBackwardTP	-0.01439	0.00567	-2.5*
Group (Older)	0.06368	0.01426	4.5*
LogFreqNoun by Group	-0.00552	0.00385	-1.4
LogBackwardTP by Group	0.00136	0.00226	0.6

Figure 1: Boxplot illustrating the effect of age group (YA vs. OA) and BackwardTP on noun duration; TP is shown for the first (lowest) and fourth (highest) quartile.



4. DISCUSSION

We investigated whether effects of probabilistic reduction, as found in studies of conversational speech, generalised to reading aloud. This study showed that the demands and complexity of oral reading did not cancel frequency and local predictability effects.

We further investigated whether frequency and transitional probability effects for noun-verb combinations change across adult age. BackwardTP indicated the predictability of the noun given the following verb, and ForwardTP always was the verb's predictability given the preceding noun. While word frequency influenced probabilistic reduction in all models, effects of transitional probability were mixed. More predictable nouns were more reduced than less predictable nouns, and this effect of BackwardTP was equal in size in younger and older adults. ForwardTP, however, had no influence on acoustic reduction of the verb (note though, that an age-by-TP interaction occurred without a simple ForwardTP effect for either group). Hence, effects of probabilistic reduction were generally similar in size in younger and older adults.

Our finding that BackwardTP effects on duration are more robust than those of ForwardTP are in line with earlier results [5]. Bell and colleagues found BackwardTP effects for both content and function words in conversational speech, while ForwardTP influenced only the production of high-frequency function words. The authors attributed this finding to differential lexical access for content versus function words in speech production. Since we investigated TP effects on content words only (verbs and nouns), it is not surprising that we only found consistent effects of BackwardTP in both age groups.

The influence of frequency and TP on speech production is often explained by a coordination mechanism that links lexical access and articulation, such that the progress of lexical retrieval is synchronized with speed of articulation [5]. Our findings of equal probabilistic reduction effects for younger and older adults suggest that the coordination between those levels is well maintained in later adulthood.

Furthermore, the results of equal TP effects for older and younger adults suggest that local TP effects differ from higher-level semantic prediction as frequently measured by cloze or sentence completion tasks [6, 18], in which older adults usually show smaller predictability effects than younger adults. Clearly, more research is needed to distinguish between TP effects and cloze predictability effects and to explore their possible change across the lifespan.

5. REFERENCES

- [1] Ashby, J., Yang, J., Evans, K. C., Rayner, K. 2012. Eye movements and the perceptual span in silent and oral reading. *Attention, Perception, & Psychophysics*, 74(4), 634-640.
- [2] Baker, R. E., Bradlow, A. R. 2009. Variability in Word Duration as a Function of Probability, Speech Style, and Prosody. *Language and Speech* 52, 391-413.
- [3] Barr, D. J., Levy, R., Scheepers, C., Tily, H. J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255-278.
- [4] Bates, D., Maechler, M., Bolker, B. 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. <http://lme4.r-forge.r-project.org/>
- [5] Bell, A., Brenier, J. M., Gregory, M., Girand, C., Jurafsky, D. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60, 92-111.
- [6] Federmeier, K. D., Kutas, M., Schul, R. 2010. Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language* 115, 149-161.
- [7] Frisson, S., Rayner, K., Pickering, M. J. 2005. Effects of Contextual Predictability and Transitional Probability on Eye Movements during Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31, 862-877.
- [8] Gollan, T. H., Montoya, R. I., Cera, C., Sandoval, T. C. 2008. More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language* 58, 787-814.
- [9] Hanique, I., Ernestus, M. 2011. Final /t/ reduction in Dutch past-participles: The role of word predictability and morphological decomposability. *Proceedings of 12th Interspeech* Florence, 2849-2852.
- [10] Inhoff, A. W., Solomon, M., Radach, R., Seymour, B. A. 2011. Temporal dynamics of the eye-voice span and eye movement control during oral reading. *Journal of Cognitive Psychology* 23, 543-558.
- [11] Jurafsky, D., Bell, A., Gregory, M., Raymond, W. D. 2001. Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In: Bybee, J. L., Hopper, P. J. (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 229 - 254.
- [12] Keuleers, E., Brysbaert, M., New, B. 2010. SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods* 42, 643-650.
- [13] McDonald, S. A., Shillcock, R. C. 2003. Eye Movements Reveal the On-Line Computation of Lexical Probabilities During Reading. *Psychological Science* 14, 648-652.
- [14] Pluymaekers, M., Ernestus, M., Baayen, R. H. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *The Journal of the Acoustical Society of America* 118, 2561-2569.
- [15] Praat: Doing Phonetics by Computer, Boersma, P., Weenink, D., 2014. <http://www.fon.hum.uva.nl/praat/>
- [16] R Development Core Team 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>
- [17] Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., Pollatsek, A. 2006. The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging* 21, 448-465.
- [18] Smith, N. J., Levy, R. 2011. Cloze but no cigar: The complex relationship between cloze, corpus, and subjective probabilities in language processing. *Proceedings of 33rd CogSci* Austin, 1637-1642.
- [19] Wurm, L. H., Fisičaro, S. A. 2014. What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language* 72, 37-48.