Aging, Neuropsychology, and Cognition
A Journal on Normal and Dysfunctional Development

ISSN: 1382-5585 (Print) 1744-4128 (Online) Journal homepage: https://www.tandfonline.com/loi/nanc20

Agreement processing and attraction errors in aging: evidence from subject-verb agreement in German

Jana Reifegerste, Franziska Hauer & Claudia Felser

To cite this article: Jana Reifegerste, Franziska Hauer & Claudia Felser (2017) Agreement processing and attraction errors in aging: evidence from subject-verb agreement in German, Aging, Neuropsychology, and Cognition, 24:6, 672-702, DOI: 10.1080/13825585.2016.1251550

To link to this article: https://doi.org/10.1080/13825585.2016.1251550

Published online: 07 Nov 2016.

AGING, NEUROPSYCHOLOGY, AND COGNITION, 2017
VOL. 24, NO. 6, 672–702
https://doi.org/10.1080/13825585.2016.1251550

Routledge
Taylor & Francis Group

Check for updates

Agreement processing and attraction errors in aging: evidence from subject-verb agreement in German

Jana Reifegerste, Franziska Hauer and Claudia Felser

Potsdam Research Institute for Multilingualism, University of Potsdam, Potsdam, Germany

ABSTRACT

Effects of aging on lexical processing are well attested, but the picture is less clear for grammatical processing. Where age differences emerge, these are usually ascribed to working-memory (WM) decline. Previous studies on the influence of WM on agreement computation have yielded inconclusive results, and work on aging and subject-verb agreement processing is lacking. In two experiments (Experiment 1: timed grammaticality judgment, Experiment 2: self-paced reading + WM test), we investigated older (OA) and younger (YA) adults' susceptibility to agreement attraction errors. We found longer reading latencies and judgment reaction times (RTs) for OAs. Further, OAs, particularly those with low WM scores, were more accepting of sentences with attraction errors than YAs. OAs showed longer reading latencies for ungrammatical sentences, again modulated by WM, than YAs. Our results indicate that OAs have greater difficulty blocking intervening nouns from interfering with the computation of agreement dependencies. WM can modulate this effect.

ARTICLE HISTORY
Received 24 May 2016
Accepted 15 October 2016

KEYWORDS
Subject-verb agreement; attraction errors; aging; grammaticality judgment; self-paced reading

Aging is associated with a decline in performance in tasks measuring general processing speed (Salthouse, 1996, 2000; Verhaeghen & Salthouse, 1997), working memory (WM) (Carpenter, Miyake, & Just, 1994; Dobbs & Rule, 1989; Foos, 1989; Just & Carpenter, 1992; Kemtes & Kemper, 1997; Light & Anderson, 1985; Verhaeghen, Marcoen, & Goossens, 1993), and long-term declarative memory learning (Park et al., 2002; Prull, Gabrieli, & Bunge, 2000) and retrieval (Craik & McDowd, 1987; Rabinowitz, 1984; Smith, 1977). Procedural memory, on the other hand, has been argued to be relatively spared from age-related cognitive decline, at least with regard to simple structures and sequences (Churchill, Stanis, Press, Kushelev, & Greenough, 2003; Frensch & Miner, 1994; Howard & Howard, 1989, 1992). Note, however, that this claim is not uncontested, and that more recent studies have established a complexity threshold for sequence learning above which age differences emerge (Bennett, Howard, & Howard, 2007; Curran, 1997; Feeney, Howard, & Howard, 2002; Howard, Howard, Dennis, Yankovich, & Vaidya, 2004; Howard, Howard, Japikse et al., 2004; Prull et al., 2000).

This dissociation in the degree to which declarative and procedural memory resources are affected by aging appears to extend to the language domain. On the

CONTACT Jana Reifegerste jana.reifegerste@uni-potsdam.de

one hand, older people show a decline in tasks relating to lexical access. They are slower and make more mistakes during picture naming (see Feyereisen, 1997; for a meta-analysis; see Burke & Shafto, 2008; for a review), they have greater word finding difficulties (Meinzer et al., 2012a, 2012b; Schmitter-Edgecombe, Vesneski, & Jones, 2000), they show a higher rate of spontaneous and induced tip-of-tongue incidents (Burke, MacKay, Worthley, & Wade, 1991; Cross & Burke, 2004; Heine, Ober, & Shenaut, 1999; Rastle & Burke, 1996; Shafto, Burke, Stamatakis, Tam, & Tyler, 2007), and their speech is marked by an increase in hesitations and pauses (Kemper, Herman, & Lian, 2003).

On the other hand, grammatical processing seems to be largely spared from age-related decline (Shafto & Tyler, 2014). Using a word-monitoring task, Tyler et al. (2010) examined age effects on the word-position effect, which taps into the online construction of linguistic representations spanning the entire sentence. The authors did not find any differences in behavioral performance, despite grey-matter loss in their older participants. Similarly, Davis, Zhuang, Wright, and Tyler (2014) report no global or specific effects of age on syntactic ambiguity resolution. In a crossmodal priming study, Clahsen and Reifegerste (in press) investigated the effects of aging on the morphological processing of participles. Besides a global effect of age (longer naming latencies for older people), they also found specific effects of age on priming for different participle types. While younger people showed significant morphological priming for irregular forms (which are stored lexical items), older speakers did not show any priming for these forms. For regular participles (which are the result of grammatical computations), however, both age groups showed priming effects of similar magnitude. The authors interpreted these findings as a result of an age-related weakening of links between stored lexical items, while grammatical processing may be preserved in older age.

Interestingly, as with effects of aging on tasks involving procedural memory, some studies point toward a complexity threshold for syntactic processing above which age effects emerge, especially with regard to language production. Kemper (1986) investigated age differences in the ability to imitate sentences of varying complexity. She found that older speakers made more errors than younger speakers imitating sentences with longer (compared to shorter) embeddings and sentences with a sentence-initial (compared to a sentence-final) embedding. For sentences with shorter sentence-final embeddings, older speakers' performance did not differ from that of younger speakers. The author attributed these age effects to increased processing demands in more complex sentences.

The present paper seeks to establish whether there are age differences in the way younger and older people process subject-verb agreement in German and the degree to which age affects the susceptibility to agreement attraction errors (Bock & Miller, 1991).

Subject-verb agreement and attraction errors

In the majority of the world's languages, structurally dependent sentence constituents agree with each other in one or more of their morphosyntactic features (Mallinson & Blake, 1981). In German, the language under study in the present paper, the subject of a sentence and the finite verb obligatorily agree in person (1st/2nd/3rd) and number (singular/plural).

As in most languages, subject-verb agreement in German is acquired early (around the age of 2; Poeppel & Wexler, 1993), and considering the staggering amount of occurrences (at least once every 5 seconds; Bock, 2003), its computation is normally relatively error-free. However, there are circumstances that increase the likelihood of errors arising, one of the most common of which is agreement attraction.

In agreement attraction, the verb agrees with a noun other than the subject. In the sentence **The key to the cabinets were rusty*, the verb *were* agrees with the noun *cabinets* (the attractor or local noun) rather than the subject of the sentence *key*. This phenomenon was initially analyzed as a kind of proximity concord (Quirk, Greenbaum, Leech, & Svartvik, 1972), in which a verb agrees with a more proximal noun phrase (NP) rather than the subject noun phrase (NP), as this is the most common triggering environment for attraction errors. It has been attested in a number of studies for various languages, and for both production and comprehension (English: Bock & Cutting, 1992; Bock & Eberhard, 1993; Bock & Miller, 1991; Pearlmutter, Garnsey, & Bock, 1999; Arabic: Tucker, Idrissi, & Almeida, 2015; Dutch: Hartsuiker, Antón-Méndez, & Van Zee, 2001; Hartsuiker, Schriefers, Bock, & Kikstra, 2003; French: Vigliocco, Hartsuiker, Jarema, & Kolk, 1996; German: Hartsuiker et al., 2003; Häussler, 2009; Hebrew: Deutsch & Dank, 2009; Italian: Vigliocco, Butterworth, & Semenza, 1995; Russian: Lorimor, Bock, Zalkind, Sheyman, & Beard, 2008; Spanish: Vigliocco, Butterworth, & Garrett, 1996; Lago, Shalom, Sigman, Lau, & Phillips, 2014).

Interestingly, attraction errors have also been found in cases where the attractor noun does not encroach on the agreement dependency between the subject and the verb. That is, in sentences such as **The drivers who the runner wave to honk back*, participants produce as well as fail to notice the number mismatch between the embedded subject NP *the runner* and the verb *wave* due to the plural feature of the main clause subject *the drivers* (Clifton, Frazier, & Deevy, 1999; Kimball & Aissen, 1971; Wagers, Lau, & Phillips, 2009). This finding is of particular importance when considering potential causes for the emergence of attraction errors. Early theories attributed these errors to faulty representations of the subject NP due to feature movement or "percolation" (Eberhard, Cutting, & Bock, 2005; Franck, Vigliocco, & Nicol, 2002; Nicol, Forster, & Veres, 1997; Vigliocco & Nicol, 1998). The idea is that feature information (such as number) can be transmitted between constituents, especially if these are structurally close. In other words, in some cases the plural number feature of *cabinets* percolates upwards to the NP node headed by *key*, thereby "overwriting" the singular number feature of the subject NP (Pearlmutter et al., 1999). Findings such as those by Wagers et al. (2009), on the other hand, challenge this explanation and instead support accounts that consider difficulties in cue retrieval as the basis of attraction errors. Under this account, both the subject NP and the attractor NP are simultaneously activated during sentence planning and comprehension. Attraction errors emerge as a consequence of faulty retrieval of the correct NP and its number feature. This faulty-retrieval account is further supported by a "grammatical asymmetry" (Wagers et al., 2009). Attraction errors occur almost exclusively in ungrammatical sentences, while grammatical sentences such as *The key to the cabinets is rusty* rarely lead to incorrect rejections from participants. This asymmetry is difficult to explain under feature-percolation accounts, which would predict the plural feature of the attractor to percolate upwards in grammatical sentences as well.

Factors influencing agreement attraction

Since the seminal work by Bock and Miller (1991), much attention has been devoted to establishing factors that influence the extent to which attraction errors occur. On an item-level, it has been proposed that the morphosyntactic number status of the attractor noun (rather than phonological features) matters (Bock & Eberhard, 1993). Pseudo-plurals such as *rose* have been argued to behave like singular forms such as *row*, rather than like its homophone counterpart *rows*. Note, however, that this conclusion has been challenged (Franck, Vigliocco, Antón-Méndez, Collina, & Frauenfelder, 2008; Hartsuiker et al., 2003; Haskell & MacDonald, 2003). Further, the notional number or numerosity of the referent of the subject plays a role during agreement processing (Eberhard, 1999; Hartsuiker, Kolk, & Huinck, 1999; Humphreys & Bock, 2005; Vigliocco, Hartsuiker et al., 1996; Vigliocco et al., 1995). So-called "distributive-referent" items (e.g., **The label of the bottles were interesting*), in which the singular grammatical number of the subject does not match with its plural conceptual number, elicit more attraction errors than single-referent items (e.g., **The author of the books were famous*), where the subject is both grammatically and conceptually singular.

On the other hand, subject-level factors that influence agreement processing have received much less attention. Kavé and Levy (2005) investigated the effect of gender congruency in subject-verb agreement in Hebrew, a language in which subjects and verbs obligatorily agree in gender, number, and person. Previous research had found longer reading latencies for verbs that mismatched the subject in number or gender (Deutsch & Bentin, 1994). Comparing younger and older participants on this phenomenon, Kavé and Levy (2005) replicated the congruency effect (longer reaction times [RTs] for incongruent contexts than for congruent contexts) and further found a main effect of age (longer RTs for older participants compared to younger participants). Importantly, congruency interacted with age: older people needed disproportionally more time to process incongruent items. The authors considered general age-related slowing as a source of this age difference, as the effects of more challenging experimental manipulations (here: gender congruency) on RTs are more influenced by general slowing than those of less challenging factors (see Laver & Burke, 1993). Further, Kavé and Levy point to a decline in executive functioning, especially inhibition, as another potential cause for the interaction, although they did not measure inhibitory skills in their participants.

In the present study, we ask whether younger and older people differ in their agreement-processing abilities. To the best of our knowledge, there is no published work that focuses on the effects of aging on the computation of subject-verb number agreement. In Experiment 1, we used a timed grammaticality-judgment task to assess the degree to which age affects the acceptance of sentences containing agreement errors. In Experiment 2, we examined younger and older people's susceptibility to agreement attraction errors using a self-paced reading task.

Experiment 1

Experiment 1 compared younger and older native speakers of German on their speed and accuracy of judging the grammaticality of German sentences with and without agreement errors. Keeping both the number of the subject (singular) and the number of the

Table 1. Overview of the stimulus materials for Experiment 1.

		Verb number	
		Singular	Plural
Distance	Short	*Der Brief an die Freunde wurde heute abgeschickt.* The letter to the friends was sent today.	**Der Brief an die Freunde wurden heute abgeschickt.* *The letter to the friends were sent today.
	Long	*Der Brief an die lustigen Freunde von Kevin wurde heute abgeschickt.* The letter to the funny friends of Kevin was sent today.	**Der Brief an die lustigen Freunde von Kevin wurden heute abgeschickt.* *The letter to the funny friends of Kevin were sent today.

intervening noun phrase (plural) constant, we manipulated the number of the verb (singular vs. plural), resulting in matching and mismatching conditions, as shown in Table 1.

Attraction effects should be reflected in a reduced ability to reject ungrammatical sentences in which the subject and the verb mismatch in number (i.e., sentences with a plural verb) in comparison to participants' ability to correctly accept grammatical sentences (i.e., sentences containing a singular verb).

We further manipulated the distance between the subject and the verb. Already in 1924, Jespersen proposed that processing of the number feature of a verb should be more difficult the further it is away from the subject due to the "mental energy" necessary (Jespersen, 1924, *p*. 345). Empirical findings on the effect of length are mixed (Hartsuiker et al., 2001; Nicol et al., 1997; Vigliocco & Nicol, 1998; but see Bock & Cutting, 1992; Fayol, Largy, & Lemaire, 1994; Kaan, 2002).

We compared the performance of two age groups: a student population and a group of older people (minimum age = 50 years). Three possible outcomes seem plausible. First, age might not affect the results at all (see Tyler et al., 2010). Second, we might find a global effect of age. While a global effect of age on accuracy rates was not expected, age-related slowing may lead to longer RTs for the older group across all conditions. Lastly, we may also find specific age effects, that is, differences in the accuracy and RT patterns for younger and older people. On the one hand, it is possible that age interacts with the verb-number manipulation. If older people have limited executive functioning (which is needed to keep information about the subject and its number feature in memory), they might be less likely to detect an ungrammatical plural form following a plural intervening noun, leading to a greater effect of verb number on accuracy rates for older people compared to younger people. On the other hand, we may find an interaction between distance and age. As older people's processing abilities are subject to decline, processes that draw on cognitive capacities might be compromised. Kemper (1986) showed that older people were less able than younger people to imitate long constructions involving embedded gerunds and relative clauses, although they performed equally well on sentences with shorter embeddings. We may thus find that effects of subject-verb distance are stronger for older people.

Method

Participants

We recruited 34 younger (6 male, $M_{Age} = 23.7$, $SD = 4.2$, range: 18–30) and 34 older (10 male, $M_{Age} = 59.3$, $SD = 9.7$, range: 50–81) native speakers of German. Participants were recruited through online ads or contacted by email after having participatedcase (e.g., a noun in Dative c in previous experiments. All participants were right-handed, gave written informed consent prior to participation, and reported normal or corrected-to-normal vision and hearing and no neurological or language-related impairments. Participants were either paid for their participation or received university course credit.

Materials

We constructed 48 experimental and 48 filler sentences. The experimental sentences consisted of a subject NP comprising a singular noun phrase modified by a prepositional phrase (PP) which contained a plural intervening noun. The subject NP was followed by a predicate phrase containing either a singular or plural verb. The experiment had a $2 \times 2 \times 2$ design with verb number (singular/plural) and distance (short/long) as within-subjects factors (as illustrated by the example item set in Table 1) and age (younger/older) as a between-subjects factor.

For the long conditions, the intervening noun in the PP was preceded by an adjective and followed by an additional PP (usually the name of a city or a country). That is, in the long condition, the distance between subject noun and verb was three words longer.

All intervening nouns were either neuter or masculine, so the noun number was also visible in the article (*die/den*). Further, all intervening nouns in the present experiment contained an overt plural suffix.

The 48 filler items were matched for sentence length with the experimental items (8 vs. 11 words). Half of the fillers were grammatical, with the majority of these (16 out of 24) containing plural subjects and verbs so as to avoid biasing participants toward responding "no" to sentences containing a plural verb. The ungrammatical filler sentences contained grammatical violations such as a wrong case (e.g., a noun in Dative case following a preposition requiring Accusative case), a mismatch in adjective-noun case agreement, or overregularizations for irregular past-tense forms.

The stimulus materials were distributed over four lists of 96 items (48 experimental items, 48 filler items) each using a Latin Square design; that is, each item appeared in a given list in only one condition. Participants were randomly assigned to one list. None of the participants reported noticing the manipulation or using a particular strategy when answering the judgment questions.

Apparatus and procedure

The experiment was programmed and presented on a laptop computer with a 15-inch screen using DMDX (Forster & Forster, 2003). The sentences were presented word by word at a rate of 750 ms per word, with words appearing in the center of the screen in font size 36 in black letters against a white background.

Participants were tested individually in a quiet room. Before the experiment, they filled out a short biographic questionnaire. For the experiment, participants were instructed to read the words appearing on the screen silently. After the presentation of the sentence-final word (indicated by a period after the word), three question marks appeared on the screen, prompting the participant to decide whether the sentence was grammatical or ungrammatical. Responses were recorded by button press on a game-pad with the dominant hand controlling the "yes" button. If the participant did not respond within 5 s, the next sentence appeared. There was no feedback on accuracy.

The experiment was preceded by eight practice sentences. The entire experiment lasted between 15 and 20 min.

The dependent measures were RT (as measured from the onset of the question marks on the screen) and accuracy. We calculated linear mixed-effects models for the RTs and generalized linear mixed-effects models (binomial family) for the accuracy rates, using the languageR package (Baayen, 2013) and the lme4 package (Bates, Maechler, Bolker, & Walker, 2014; R Development Core Team, 2011). Using backwards elimination, we established the model that best accounted for accuracy rates and log-transformed RTs on the basis of the independent factors of items (verb number, distance), participants (age group), and trials (trial number; i.e., number of the trial within the experimental session) with participants and items as random factors. The fixed factors were centered. Following Barr, Levy, Scheepers, and Tily (2013), we started with a maximal random-factor structure and simplified the model in cases of convergence failure.

Results

We excluded trials with RTs shorter or longer than 2.5 SDs from the mean on a per-participant basis, resulting in 4.3% (younger group: 3.4%; older group: 5.2%) data loss. For the analysis of RTs, we further excluded all trials with incorrect end-of-sentence grammaticality judgments.

Accuracy

Overall, both groups of participants showed a similar pattern in that they were more accurate at accepting correct (SPS) sentences than at rejecting incorrect (*SPP) sentences. The younger participants accepted correct sentences 95.9% of the time while rejecting incorrect ones 87.8% of the time, with an overall accuracy of 91.9%. The older participants showed 94.4% accuracy for correct sentences but only 71.5% accuracy for incorrect ones, with an overall accuracy of 83.1%. Figure 1 illustrates end-of-sentence accuracy rates as a function of verb number, distance, and age group.

Verb number significantly influenced accuracy rates, with singular verbs leading to greater accuracy than plural verbs for all participants. A significant interaction between age and verb number reflected the fact that this effect was stronger for older people ($z = 3.39$, $p < .001$) than for younger people ($z = 2.50$, $p = .013$). Further, there was a main effect of distance; for both age groups, sentences with a greater distance between subject and verb had lower accuracy rates than sentences with a shorter distance (85.3% vs. 89.7%). There were no interactions of distance with either verb number ($z = -0.41$, *ns*) or age group ($t = -0.07$, *ns*). See Table 2 for the entire model.

In order to ensure that our results are not skewed by any potential group differences in response bias, we compared *A*′ scores (which take into account any response bias) between the two groups. The results showed a main effect of age group ($F = 12.27$, $p < .001$), with lower *A*′ scores for older speakers ($M = .888$, $SD = .113$) compared to younger speakers ($M = .955$; $SD = .043$). There were no effects of distance ($F = 1.98$, *ns*) and no interaction between age group and distance ($F < 1$, *ns*) on *A*′ scores.[1]

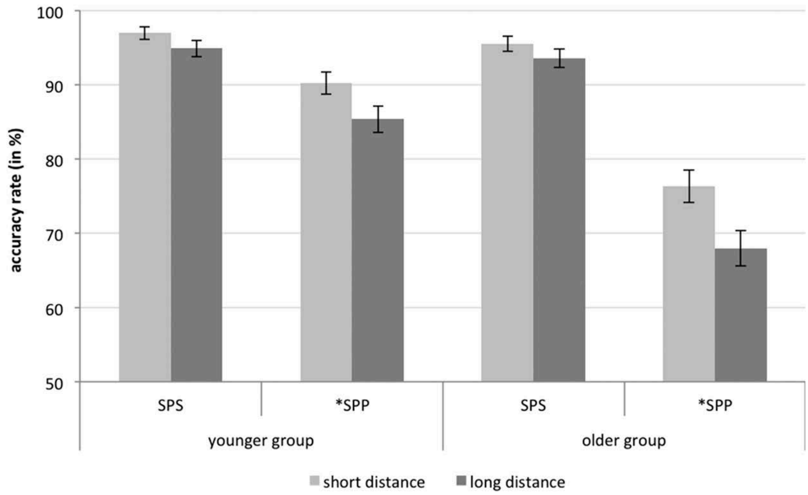

Figure 1. Judgment accuracy as a function of verb number, distance, and age group in Experiment 1. (SPS: singular subject, plural attractor, singular verb; *SPP: singular subject, plural attractor, plural verb; in all graphs, error bars reflect the standard error of the mean).

Table 2. Results of Experiment 1, linear mixed-effects model for accuracy rates.

Fixed effects		β	Standard error	*z*-value	*p*-value
Intercept		2.88	0.21	13.51	<.001
Verb number		1.96	0.16	12.37	<.001
Distance		0.56	0.13	4.25	<.001
Age group		0.69	0.33	2.12	.034
Verb number: age group		−1.09	0.32	−3.45	<.001
Verb number: distance		−0.10	0.29	−0.34	.730
Distance: age group		0.05	0.29	0.17	.867
Verb number: distance: age group		0.28	0.59	0.47	.640
Random effects		Variance	Standard deviation		Correlation
Items	Intercept	0.869	0.932		
	Age group	0.194	0.441		−0.33
Subjects	Intercept	1.088	1.043		

Formula of the best-fit model: DV ~ verb.number * age.group + distance + (1+ age.group|item) + (1|subject).

Reaction times

Figure 2 shows participants' mean RTs for end-of-sentence judgments as a function of verb number, distance, and age group.

We found a significant main effect of age group, with older people having longer RTs compared to younger people (786 ms vs. 483 ms). Further, a main effect of distance indicates that a greater distance between subject and verb led to shorter judgment times (609 ms vs. 640 ms). See Table 3 for the entire model.

An analysis of the filler items revealed the same effects as for the experiment items. Older people showed significantly longer RTs than younger people (t = −5.83), and longer sentences led to shorter RTs than shorter sentences (t = 1.97). Verb number did not affect RTs (t < 1, *ns*).

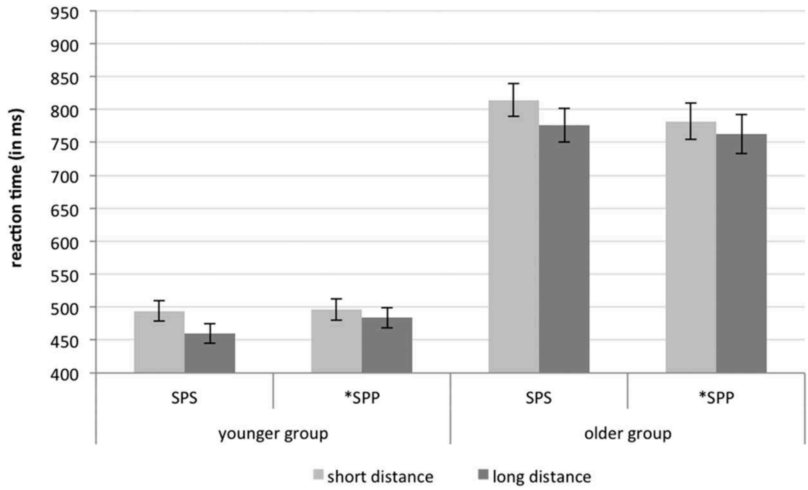

Figure 2. RTs for end-of-sentence grammaticality judgments as a function of verb number, distance, and age group in Experiment 1. (SPS: singular subject, plural attractor, singular verb; *SPP: singular subject, plural attractor, plural verb).

Table 3. Results of Experiment 1, linear mixed-effects model for RTs.

Fixed effects		β	Standard error	*t*-value
Intercept		6.27	0.05	137.96
Verb number		−0.02	0.02	−0.87
Distance		0.08	0.02	4.01
Age group		−0.51	0.09	−5.89
Verb number: age group		−0.03	0.04	−0.76
Verb number: distance		0.03	0.04	0.66
Age group: distance		−0.04	0.04	−0.98
Verb number: age group: distance		0.10	0.08	1.23
Random effects		**Variance**	**Standard deviation**	**Correlation**
Items	Intercept	0.01	0.10	
	Age group	0.01	0.05	1.00
Subjects	Intercept	0.12	0.34	
Residual		0.30	0.54	

Formula of the best-fit model: DV ~ age.group + distance + (1+ age.group|item) + (1|subject).

Discussion

The aim of Experiment 1 was to investigate subject-verb agreement computation in older people using a timed grammaticality-judgment task and to compare their performance to that of a younger control group.

In the accuracy data, we found that all participants made more errors for (incorrect) plural than for (correct) singular verbs; that is, they incorrectly accepted sentences containing a plural verb following a plural NP even though the subject NP was singular. Importantly, this effect was significantly larger for older people, whose accuracy dropped from 94% for singular to 72% for plural verbs (compared to a drop from 96% to 88% for younger people). No such effect of verb number was visible in the filler items, however (see note 1). We propose that the increased incorrect acceptance of plural verbs in our experimental items may be due to the intervening plural NP, with older people having a harder time blocking the plural NP from interfering with the correct computation of the agreement dependency between the singular subject NP and the verb. This age effect might be due to an age-related decline in executive functions, in particular cognitive control.

Additionally, we found that greater linear distance between subject and verb caused accuracy rates to decrease as well. However, this effect was similar in size for singular and plural verbs; that is, it did not lead to any further increase in the acceptance of ungrammatical sentences. This finding is in line with previous work that did not find any specific effects of linear distance on attraction errors (e.g., Hartsuiker et al., 2001; Vigliocco & Nicol, 1998). In our study, one possible reason for this is that the distance manipulation did not only create greater distance between the subject NP and the verb but also between the attractor NP and the verb. For example, in the long condition **Der Brief an die lustigen Freunde von Kevin wurden …*, the plural verb *wurden* is three words further away from the singular subject NP *der Brief* compared to its short equivalent **Der Brief an die Freunde wurden …*, but also two words further away from the plural attractor NP *die Freunde*. The observed effect of distance may thus simply reflect between-condition differences in complexity (as the PP following the attractor noun made the noun phrase more complex) or of sentence length (as our long sentences were three words longer than the short ones; see Martin & Roberts, 1967; Wearing, 1969).

The distance manipulation did not affect older people to a greater extent than younger ones. While we found a numerical difference between the drop in accuracy for shorter vs. longer sentences for singular verbs (younger and older: 2% drop) compared to plural verbs (younger: 5% drop, older: 8% drop), the relevant three-way interaction between verb number, distance, and age group did not prove significant.

Regarding participants' end-of-sentence judgment times, we found that older people generally responded more slowly than younger people, an effect that can be attributed to general age-related slowing (Salthouse, 1996, 2000).

We also found that long sentences elicited shorter RTs than short sentences. Importantly, an analysis of our fillers yielded the same effect, suggesting that this is not an effect of subject-verb distance but an effect of sentence length or complexity. We speculate that this length effect might reflect some kind of trade-off between online processing effort and sentence-final judgment times. Our longer sentences were likely to be more difficult to process than the shorter ones (as evidenced by lower accuracy scores), so that the additional mental effort required to process the former might have facilitated grammaticality judgments.

Unlike in the accuracy data, we did not find a main effect of verb number or an interaction between age group and verb number for RTs. One might have expected such effects to arise; specifically, sentences containing ungrammatical plural verbs might be expected to elicit longer RTs than sentences containing grammatical singular verbs. One explanation for the absence of this effect in our results could be that participants made the grammatical judgment before encountering the judgment question (i.e., while still reading the sentence).

To summarize, in the present experiment we found that older people had a reduced ability, in comparison to younger people, to correctly identify ungrammatical sentences that contained a plural verb following a plural intervening NP. While we found a global effect of age on response latencies (i.e., longer RTs for older people), this effect was not disproportionally larger when speakers had to reject an incorrect sentence with a plural verb. One plausible explanation for the difference in accuracy rates between younger and older people are age-related differences in executive functioning, especially with regard to WM capacities. Processing subject-verb agreement involves holding the subject in WM, so the parser can verify the required feature matching once it encounters the verb. Declining WM abilities may lead to an increased likelihood of errors. Experiment 2 will follow-up on this explanation.

To conclude with an important caveat, note that our experimental design confounds verb number with agreement attraction. Although we attribute older speakers' lower accuracy rates in rejecting incorrect plural verbs to interference from intervening plural NPs, agreement attraction is not the only possibility that may have led to this result. Note that German plural forms are one or two letters longer than their corresponding singular forms and of lower frequency, which may have made them more difficult to process especially for our older speakers. It is also conceivable that plural forms are conceptually more difficult to process than singular forms, which may again have affected older speakers to a greater extent than younger speakers. Experiment 2 was designed so as to provide a more direct measure of agreement attraction than Experiment 1.

Experiment 2

Having found differences in the accuracy with which younger and older speakers judged sentences containing agreement errors, we sought to investigate potential reasons for this difference. Aging by itself is not a particularly informative variable as a number of cognitive processes decline with age (as laid out in the Introduction). One of these processes, WM, has received some attention in the agreement-attraction literature.

Hartsuiker and Barkhuysen (2006) investigated the influence of WM on agreement processing. The authors manipulated extrinsic WM load by having participants maintain a word list during a sentence-completion task. They compared the incidence of attraction errors for single-referent items and distributive-referent items in young native speakers of Dutch with different levels of verbal WM. The authors found an interaction between the load manipulation and participants' verbal WM scores. Participants with low WM scores were influenced by the load manipulation and made more agreement errors, while participants with high scores were not. Notional number affected both groups of participants to a similar degree. The authors took this absence of a specific effect of notional number to mean that limited WM capacities do not interfere with the conceptual representation of the subject of a sentence (in which case notional number should have had a greater effect on speakers with low WM scores). Instead, Hartsuiker and Barkhuysen concluded that limited resources affect the speaker's ability to resolve the conflict that arises between the different NPs as both compete for specifying the number of the subject NP.

Fayol et al. (1994) found similar effects when they taxed WM with a dual-task. Participants who were asked to write down orally presented sentences made significantly more number agreement errors when they had to remember a sequence of words at the same time compared to when they could concentrate on the writing task alone (see also Hupet, Fayol, & Schelstraete, 1998). In summary, it seems that WM is involved in agreement processing. Given these findings, we conducted a second experiment that included a measure of WM capacities of our participants to explore whether the age differences we found in Experiment 1 might indeed have been due to differences in WM between younger and older speakers.

Experiment 2 was designed to extend the findings from Experiment 1 in three crucial ways. First, as mentioned above, we manipulated the number of the attractor noun. This will provide us with a baseline condition, which allows for a more direct measure of agreement attraction by comparing the erroneous acceptance of plural verbs following singular intervening NPs (*SSP) to that of plural verbs following plural intervening NPs (which we from now on refer to as "attractor NP," *SPP).

Second, following up on previous research suggesting the involvement of WM in agreement attraction (Fayol et al., 1994; Hartsuiker & Barkhuysen, 2006; Hupet et al., 1998), we measured participants' WM skills using a reading-span task based on Waters and Caplan (1996).

Third, instead of the timed word-by-word presentation used in Experiment 1, stimuli were presented using the self-paced reading paradigm before the grammaticality-judgment prompt in Experiment 2. This allowed participants to take as much time as necessary to read each word of the sentence, limiting the possibility that incorrect responses are a consequence of participants (particularly those in the older group) not being able to finish reading the words. This paradigm further allowed us to compare attraction effects during online sentence comprehension and end-of-sentence judgments in younger and older people. Some previous studies have found age differences to be present for some tasks but absent for others. Caplan and Waters (2005) compared younger and older speakers' sentence processing abilities by combining the auditory moving windows paradigm (i.e., "self-paced listening") with an end-of-sentence plausibility judgment of sentences of differing syntactic complexity. The authors found a global effect of age (i.e., longer RTs for older people) for both types of tasks. Specific age effects – that is, disproportional age-related slowing for a particularly difficult condition – were found only for the end-of-sentence judgment as well as for one out of three online-processing comparisons. Similar effects were found for correlations between WM capacity (as measured by the reading span test in which older people achieved significantly lower scores) and performance in online processing vs. end-of-sentence judgments. Caplan and Waters (2005) concluded that aging and related changes in verbal WM do not affect the efficiency with which speakers process syntactic constructions online. Offline end-of-sentence judgments, which involve metalinguistic judgments and associated review processes and decision heuristics, however, seem to be influenced by WM capacities. In Experiment 2, examining participants' reading times additionally allowed us to investigate whether the age differences we found for end-of-sentence judgments in Experiment 1 also hold for online reading latencies.

Method

Participants

We recruited 32 younger (10 male, M_{Age} = 24.4, *SD* = 4.6, range: 19–40) and 32 older (13 male, M_{Age} = 66.8, *SD* = 5.4, range: 61–78) native speakers of German. All participants were right-handed, gave written informed consent prior to participation, reported normal or corrected-to-normal vision and hearing and no neurological or language-related impairments. Participants were either paid for their participation or received course credit.

Working memory test

In order to assess participants' verbal WM capacity, we used a German version of the Reading Span test developed by Waters and Caplan (1996).

Materials. Participants were presented with 40 (acceptable and unacceptable) sentences that were either (a) cleft-subject sentences, (b) cleft-object sentences, (c) object-relativized sentences, or (d) subject-relativized sentences.

(1a) *Es war der Arbeiter, der mehr Geld wollte*.

"It was the worker that wanted more money."

(1b) *Es war der Polizist, den der Überfall verhinderte*.

"It was the police officer that the robbery prevented."

(1c) *Die Frau traf den Autor, der das Buch schrieb*.

"The woman met the author that wrote the novel."

(1d) *Der Koch, den die Suppe zubereitete, schmeckte gut*.

"The cook that the soup prepared tasted good."

Half of these sentences were semantically plausible (here: a and c), the other half were implausible.

Procedure. The sentences were presented in 11 blocks, with between two and six sentences per block. Participants were asked to read a sentence and then provide a plausibility judgment. After each block, they were asked to write down the last word of every sentence in the correct order of appearance.

Scoring. Participants could score 0.5 points for every correct plausibility judgment and 0.5 per correctly recalled word, for a total of 40 points.

Grammaticality judgment task

Materials. We used the same sentences as in the short condition in Experiment 1 (i.e., sentences with a plural attractor NP and a singular or plural verb) and added attractor number as a second factor, resulting in a 2 × 2 × 2 design with the factors attractor number (singular/plural) and verb number (singular/plural) as within-subjects factors and age group (younger/older) as a between-subjects factor. The mean lengths of singular and plural attractors were 6.5 and 7.8 letters (SDs: 2.2 and 2.1), respectively. The mean lengths of singular and plural verbs were 4.3 and 5.8 letters (SDs: 1.2 and 0.9), respectively. As in Experiment 1, all subject NPs were singular, rendering all plural verbs ungrammatical. See Table 4 for an example item set.

Table 4. Overview of the stimulus materials for Experiment 2.

		Verb number	
		Singular	Plural
Attractor number	Singular	*Der Brief an den Freund wurde heute abgeschickt.* The letter to the friend was sent today.	**Der Brief an den Freund wurden heute abgeschickt.* *The letter to the friend were sent today.
	Plural	*Der Brief an die Freunde wurde heute abgeschickt.* The letter to the friends was sent today.	**Der Brief an die Freunde wurden heute abgeschickt.* *The letter to the friends were sent today.

We added 96 filler sentences, half of which were grammatical. Ungrammatical filler sentences contained violations similar to those used in Experiment 1.

The stimulus materials were distributed over four lists of 144 items (48 experimental and 96 filler items) each using a Latin Square design; each item appeared in a given list in only one condition. Participants were randomly assigned to one list. None of the participants reported noticing the manipulation or using a strategy when they answered the judgment questions.

Apparatus and procedure. The experiment was programmed and presented on a laptop computer with a 15-inch screen. The sentences were presented word-by-word using the moving-window paradigm (Just, Carpenter, & Woolley, 1982) using Doug Rohde's Linger software (http://tedlab.mit.edu/~dr/Linger/) with participants controlling the presentation duration of each word via button-press. Words appeared in font size 36 in white letters against a black background.

Participants were tested individually in a quiet room. Before the experiment, they filled out a short biographic questionnaire. For the experiment, participants were instructed to read the words appearing on the screen at their own speed. After the presentation of the last word of a sentence, three question marks appeared on the screen, prompting the participant to decide whether or not the sentence was grammatical. Responses were recorded by button press on a gamepad. There was no timeout and no feedback on accuracy. The experiment was preceded by eight practice sentences. The entire experiment lasted between 20 and 30 min.

The dependent measures were accuracy and RTs for the grammaticality-judgment question (as measured from the onset of the question marks on the screen), and reading latencies per word. We calculated (generalized) linear mixed effects models in the same manner as in Experiment 1 to establish the model that best accounts for accuracy rates and log-transformed RTs. Fixed factors were age group, WM score[2], attractor number, verb number, attractor length in letters, verb length in letters, and trial number. Attractor length and verb length were included to separate effects of plural number from effects of word length. All fixed effects were centered. Participants and items were random factors.

Results

Working memory test

Table 5 provides an overview of the WM scores by age group.

WM scores were influenced by age, with the older group scoring significantly lower than the younger group ($t = 5.65$, $p < .001$).

Table 5. Overview of the working-memory scores by age group, Experiment 2.

	Younger	Older	Total
Mean	36.48	32.46	34.47
SD	2.38	3.23	3.47
Range	31–39.5	26–38.5	26–39.5

Grammaticality judgment task

We excluded trials with RTs beyond 2.5 SDs from the mean on a per-participant basis, resulting in 3.8% (younger group: 3.5%; older group: 4.1%) data loss. For the RT analysis, we further excluded all trials with incorrect end-of-sentence grammaticality judgments.

Accuracy. Figure 3 illustrates participants' accuracy rates for end-of-sentence judgments as a function of attractor number, verb number, age group and WM score.

Overall, both groups of participants showed a similar pattern in that they were similarly accurate for sentences containing a singular verb (i.e., correct sentences, SSS and SPS) and for incorrect sentences containing a plural verb after a singular attractor (*SSP). Both groups were the least accurate when it came to the critical *SPP sentences in which a plural verb followed a plural attractor. The difference in accuracy between *SSP and *SPP sentences constitutes the attraction effect, which was 8.7% for the younger group and 10.6% for older group. Within the older group, the attraction effect was particularly strong for people with low WM scores compared to people with high WM scores.

Attractor number significantly influenced response accuracy, with plural attractors leading to lower accuracy compared to singular NPs (91.9% vs. 97.4%). Similarly, a main effect of verb number indicated that sentences with plural verbs had lower accuracy rates than sentences with singular verbs (91.8% vs. 97.6%). While attractor number and verb number did not interact across all participants ($z = -0.07$, *ns*), a three-way interaction between attractor number, verb number, and age, as well as a four-way interaction

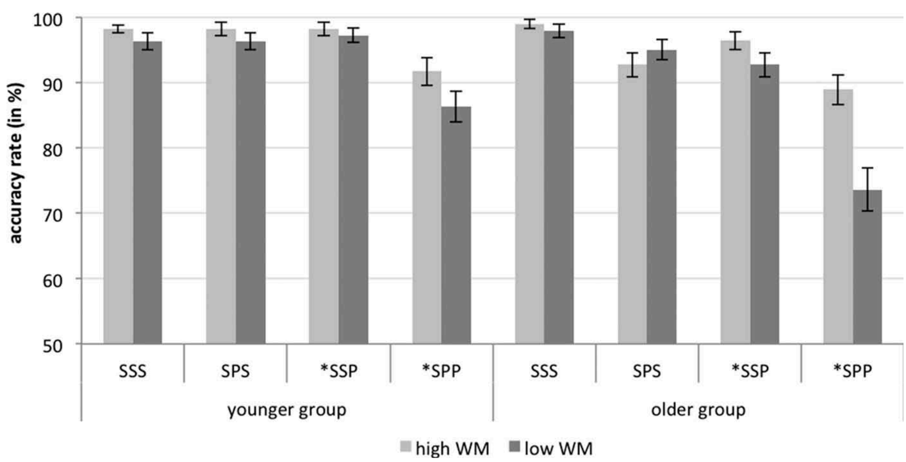

Figure 3. Accuracy of end-of-sentence grammaticality judgments as a function of verb number, distance, age group, and working-memory score in Experiment 2. (SSS: singular subject, singular attractor, singular verb; SPS: singular subject, plural attractor, singular verb; *SSP: singular subject, singular attractor, plural verb; *SPP: singular subject, plural attractor, plural verb).

Table 6. Results of Experiment 2, linear mixed-effects model for accuracy rates.

Fixed effects		β	Standard error	*z*-value	*p*-value
Intercept		3.88	0.30	13.17	<.001
Attractor number		1.57	0.43	3.67	<.001
Verb number		1.24	0.34	3.64	<.001
WM score		0.25	0.07	3.57	<.001
Age group		−0.70	0.48	−1.48	.140
Trial number		0.02	0.01	2.79	.005
Attractor number: verb number		−0.05	0.68	−0.07	.942
Attractor number: WM score		0.03	0.10	0.33	.745
Verb number: WM score		−0.02	0.10	−0.19	.847
Attractor number: age group		−0.86	0.69	−1.24	.214
Verb number: age group		−0.73	0.68	−1.07	.283
WM score: age group		0.09	0.14	0.67	.501
Attractor number: verb number: WM score		0.14	0.19	0.72	.471
Attractor number: verb number: age		−3.21	1.37	−2.34	.019
Attractor number: WM score: age group		−0.16	0.19	−0.84	.403
Verb number: WM score: age group		0.05	0.19	0.28	.777
Attractor number: verb number: age: WM score		−0.76	0.38	−2.01	.045
Random effects		Variance	Standard deviation		Correlation
Items	Intercept	0.55	0.74		
Subjects	Intercept	0.90	0.95		
	Attractor number	0.03	0.18		1.00

Formula of the best-fit model: DV ~ attractor.number * verb.number * age.group * WM + trial + (1|item) + (1+attractor.number|subject).

between attractor number, verb number, age, and WM score indicate different accuracy patterns for younger and older speakers. See Table 6 for the entire model.

In order to understand the nature of these age differences, we analyzed the two age groups separately.

Younger speakers. For younger people, accuracy rates were significantly influenced by the number of the attractor; plural attractors led to lower accuracy than singular attractors (93.6% vs. 97.8%). Similarly, a main effect of verb number indicates that sentences with plural verbs elicited significantly lower accuracy scores than those with singular verbs (93.5% vs. 97.8%). A significant interaction between attractor number and verb number revealed an attraction effect. Participants were significantly worse at rejecting sentences in which an incorrect plural verb was preceded by a plural noun (*SPP) compared to sentences where the incorrect plural verb was preceded by a singular noun (*SSP), $z = 2.13$, $p = .034$. There was no such effect for singular verbs ($z = -0.78$, *ns*). In other words, participants were more likely to accept an incorrect plural verb when it was preceded by a plural compared to a singular noun. Importantly, while a low WM score led to lower accuracy scores, this effect was present across all conditions and did not modulate any of the other effects. See Table 7 for the entire model for younger speakers.

Older speakers. For older speakers, similar to younger speakers, we found main effects of attractor number, with plural attractors leading to lower accuracy than singular nouns (90.3% vs. 97.0%), and of verb number, with plural verbs yielding lower accuracy scores than singular verbs (90.1% vs. 97.3%). Importantly, besides a main effect of WM score, we also found an interaction between attractor number, verb number, and WM score. WM score seems to affect the strength of the interaction between attractor number and verb number: The lower a participant's WM score, the stronger the interaction between attractor number and verb number. In other words, a person with a relatively low WM

Table 7. Results of Experiment 2 for younger speakers, linear mixed-effects model for accuracy rates.

Fixed effects		β	Standard error	*z*-value	*p*-value
Intercept		3.69	0.42	8.71	<.001
Attractor number		1.51	0.62	2.45	.014
Verb number		0.88	0.37	2.36	.018
WM score		0.30	0.11	−2.66	.008
Trial number		0.01	0.01	0.62	.538
Attractor number: verb number		−1.61	0.74	−2.17	.030
Attractor number: WM score		−0.02	0.18	−0.11	.915
Verb number: WM score		0.03	0.15	0.17	.867
Attractor number: verb number: WM score		−0.23	0.30	−0.76	.447
Random effects		Variance	Standard deviation		Correlation
Items	Intercept	0.83	0.91		
Subjects	Intercept	0.89	0.94		
	Attractor number	0.74	0.86		0.87

Formula of the best-fit model: DV ~ attractor.number * verb.number + WM + (1|item) + (1+attractor.number|subject).

Table 8. Results of Experiment 2 for older speakers, linear mixed-effects model for accuracy rates.

Fixed effects		β	Standard error	*z*-value	*p*-value
Intercept		4.14	0.41	10.22	<.001
Attractor number		1.82	0.64	2.83	.004
Verb number		1.50	0.56	2.69	.007
WM score		0.20	0.08	2.38	.017
Trial number		0.03	0.01	2.79	.005
Attractor number: verb number		1.51	1.12	1.34	.179
Attractor number .: WM score		0.10	0.12	0.79	.427
Verb number: WM score		−0.07	0.12	−0.60	.546
Attractor number: verb number: WM score		0.53	0.74	2.19	.029
Random effects		Variance	Standard deviation		Correlation
Items	Intercept	0.36	0.60		
Subjects	Intercept	0.89	0.94		
	Attractor number	0.01	0.02		0.64

Formula of the best-fit model: DV ~ attractor.number * verb.number * WM + trial + (1|item) + (1+attractor.number|subject).

score showed a significantly larger attraction effect than a person with a relatively high WM score. See Table 8 for the entire model for older speakers.

Grammaticality judgment times. Figure 4 illustrates participants' end-of-sentence judgment RTs as a function of attractor number, verb number, age group and WM score.

We found a significant main effect of age group (young: 654 ms, old: 901 ms) as well as a marginal main effect of verb number, with singular verbs eliciting shorter RTs than plural verbs (770 ms vs. 784 ms). Further, age group interacted with verb number; while younger people had longer RTs for sentences with (incorrect) plural verbs, older people showed longer RTs for sentences with (correct) singular verbs (t = 2.70). There were no effects of WM score on end-of-sentence judgment RTs (t = 1.33). See Table 9 for the entire model.

Self-paced reading latencies. Figure 5 illustrates reading latencies for the different sentence types broken down by age group.

See Table 10 for an overview of all significant effects that were found for self-paced reading latencies. For all words, we found a main effect of age (longer latencies for older

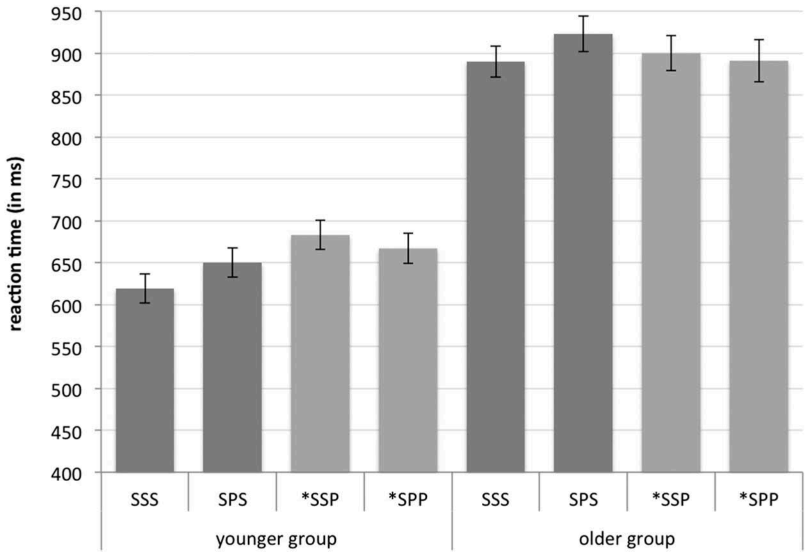

Figure 4. RTs for end-of-sentence judgments as a function of attractor number, verb number, and age group in Experiment 2. Darker bars represent grammatical sentences, lighter bars ungrammatical sentences. (SSS = singular subject, singular attractor, singular verb; SPS: singular subject, plural attractor, singular verb; *SSP: singular subject, singular attractor, plural verb; *SPP: singular subject, plural attractor, plural verb).

Table 9. Results of Experiment 2, linear mixed-effects model for grammaticality-judgment times.

Fixed effects		β	Standard error	*t*-value
Intercept		6.55	0.04	180.51
Verb number		−0.02	0.01	−1.82
Age group		−0.34	0.07	−4.66
Trial number		−0.003	0.001	−7.74
Attractor number		−0.014	0.013	−1.12
Verb number: age group		−0.10	0.03	−3.89
Attractor number: WM score		0.001	0.005	0.21
Verb number: WM score		−0.001	0.005	−0.26
Attractor number: age group		−0.023	0.032	−0.71
WM score: age group		0.010	0.028	0.36
Attractor number: verb number		−0.049	0.050	−0.97
Attractor number: WM score		0.001	0.005	0.21
Verb.number: WM score		−0.001	0.005	−0.26
Attractor number: age group		−0.023	0.032	−0.71
Verb number: WM score: age group		−0.002	0.010	−0.16
Attractor number: WM score: age group		−0.001	0.010	−0.09
Attractor number: verb number: WM score		−0.001	0.010	−0.07
Attractor number: verb number: age group		−0.050	0.064	−0.78
Attractor number: verb number: WM score: age group		0.019	0.019	0.98
Random effects		**Variance**	**Standard deviation**	**Correlation**
Items	Intercept	0.001	0.026	
Subjects	Intercept	0.081	0.285	
	Attractor number	0.001	0.007	−1.00
Residual		0.108	0.329	

Formula of the best-fit model: DV ~ verb.number * age.group + trial + (1|item) + (1+attractor.number|subject).

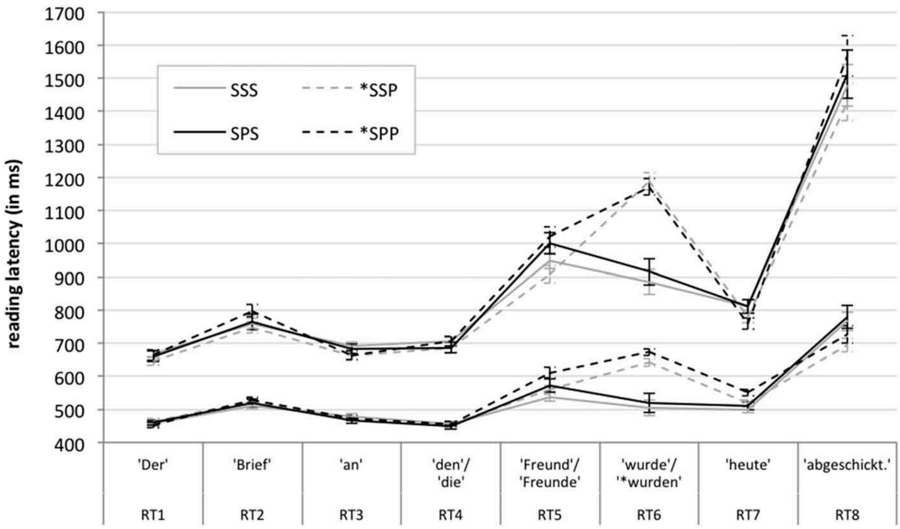

Figure 5. Word-by-word reading latencies for the two age groups (top: older group, bottom: younger group) in Experiment 2. (SSS: singular subject, singular attractor, singular verb; SPS: singular subject, plural attractor, singular verb; *SSP: singular subject, singular attractor, plural verb; *SPP: singular subject, plural attractor, plural verb).

people than for younger people) and of trial number (responses became faster over the course of the experiment). Besides these effects, there were no significant main effects or interactions on the first four words, which were identical across conditions.

On the attractor word (the fifth word), we found a main effect of attractor number, with longer latencies for plural than for singular NPs (795 ms vs. 736 ms), as well as a main effect of attractor length in letters, indicating longer reading times for longer

Table 10. Results of Experiment 2, significant effects on self-paced reading times.

Position and example	Effect	β	Standard error	*t*-value
1 *Der*	Age group	−0.353	0.061	−5.74
	Trial number	−0.002	0.0002	−8.81
2 *Brief*	Age group	−0.394	0.079	−5.00
	Trial number	−0.003	0.0004	−7.99
3 *an*	Age group	−0.354	0.064	−5.51
	Trial number	−0.003	0.0003	−8.33
4 *den/die*	Age group	−0.414	0.068	−6.13
	Trial number	−0.002	0.0003	−5.55
5 *Freund/Freunde*	Attractor number	−0.055	0.016	−3.50
	Attractor length	0.026	0.005	5.30
	Age group	−0.497	0.084	−5.92
	Trial number	−0.002	0.0004	−5.47
6 *wurde/*wurden*	Verb number	−0.081	0.019	−4.37
	Verb length	0.091	0.011	8.61
	Attractor number	−0.030	0.011	−2.64
	Age group	−0.577	0.096	−5.99
	Trial number	−0.004	0.0004	−8.65
	Age group: verb length	−0.046	0.016	−2.82
	Age group: verb number	0.078	0.033	2.38
	Age group: verb number: WM	−0.019	0.010	−1.94
7 *heute*	Age group	−0.334	0.068	−4.93
	Trial number	−0.002	0.0004	−6.42
	Age group: verb number	−0.091	0.021	−4.35

words compared to shorter words. The fact that both of these effects came out as significant indicates that longer reading times are not solely due to plural forms being longer, but are also partly explained by longer processing times of plural forms due to their plural status.

On the critical verb (the sixth word), we found several main effects and interactions. A main effect of verb number indicates that plural verbs took longer to read than singular verbs (909 ms vs. 706 ms). A main effect of verb length in letters suggests that longer forms led to longer reading times. As with the attractor nouns, the fact that both effects came out as significant suggests that the effect of verb number is not solely due to plural verb forms being longer, but instead also due to their plural status. An interaction between verb number and age group indicates that the effect of verb number was stronger for older people than for younger people (younger people: 172 ms difference, older people: 350 ms difference). A main effect of attractor number suggests that the slowing effect of plural attractors spilled over from the previous word, leading to longer reading times of the verb if the preceding noun was a plural form. An interaction between age group and verb length in letters indicated that the slowing effect of longer verbs was larger for older people. Crucially, we found a marginally significant three-way interaction between verb number, WM score, and age. The interaction between verb number and WM score holds only for older people (t = 2.77), but not for younger people (t = 0.22). In other words, the extent to which a person had longer reading times for plural verbs was modulated by WM score within older people only. See Figure 6 for an illustration of this effect on the critical word.

On the seventh word (critical word+1), we found an interaction between verb number and age. While younger people showed longer RTs for plural verbs than for singular verbs as they had on the previous word (534 ms vs. 504 ms, t = −3.05), older people showed the opposite pattern with longer latencies for singular verbs than for plural verbs (811 ms vs. 770 ms, t = 2.70); that is, the effect of verb number reversed at the next word for older people, but stayed in the same direction for younger people.

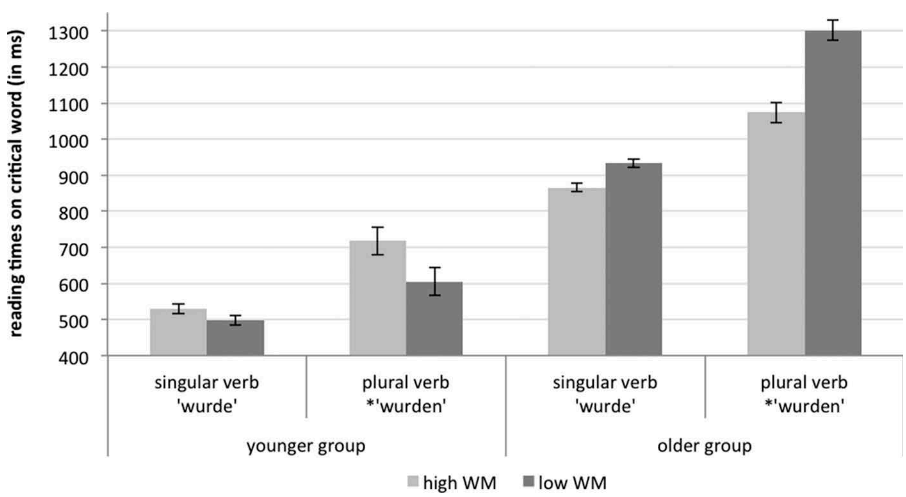

Figure 6. Reading latencies on the critical word as a function of verb number, age group, and working-memory score in Experiment 2.

We did not analyze reading latencies on the last word as those are affected by wrap-up processes and may include part of the decision component.

Discussion

Experiment 2 was designed to extend the findings from Experiment 1 in two ways. First, we combined the grammaticality-judgment task with the self-paced reading paradigm to gain further insights into the online processing of agreement in older people. Second, we included a WM measure to test whether age effects in agreement processing might be affected by age-related changes in WM skills.

Unlike in Experiment 1, we manipulated both the number of the attractor NP (singular vs. plural) and the number of the verb (singular vs. plural). All experimental subject NPs were singular, rendering all sentences containing plural verbs ungrammatical. Of particular interest was the difference between the conditions *SPP and *SSP – if a plural verb following a plural noun leads to lower accuracy scores and/or longer RTs than a plural verb following a singular noun, this is assumed to be due to number attraction from the NP preceding the verb.

With regard to end-of-sentence judgment accuracy, we found that younger and older people showed different patterns and that these patterns were influenced by WM scores for the older participants, but not for younger participants. For the younger group, all participants showed a similar degree of decline in accuracy for the most difficult condition, in which a plural verb was preceded by a plural attractor noun. Among the older group, on the other hand, WM modulated their performance in the grammaticality-judgment task. Older people with relatively low WM scores showed a rather dramatic drop in accuracy of almost 15%, whereas older people with relatively high WM scores performed as well as those in the younger group. In other words, for older people, a limited WM capacity hampered their ability to detect attraction errors, while for younger people it did not. In neither group was there a difference in accuracy rates for SSS and SPS sentences, replicating the "grammaticality asymmetry" according to which attraction occurs in ungrammatical but not in grammatical sentences (Wagers et al., 2009).

Turning to participants' reading latencies, we found that both groups were sensitive to the verb-number match/mismatch manipulation: Both younger and older people showed longer latencies when the verb did not match in number with the subject. This effect was stronger for older people, who took twice as long before continuing to the next word. Similar to the accuracy data, the size of the number effect was modulated by WM in the older but not in the younger group. In other words, for older people only, participants with a low WM score showed a larger effect of verb number than participants with a high WM score. WM did not influence younger people's reading latencies on the critical word, however.

A possible reason for this difference could be that older participants – in particular those with relatively poor WM skills – needed more time than younger people to retrieve the number information of the subject from WM, a process that is slower and more effortful for older people (Carpenter et al., 1994; Kemper, 1986; Light & Anderson, 1985). A second possibility might be that the two age groups employed different strategies during the self-paced reading task, especially in combination with end-of-sentence grammaticality judgments. While older people only continued to the next word when they had processed the current word completely (and, in the case of a plural verb, realized that the sentence was incorrect), the younger group may have moved on to the next word more quickly and made their grammaticality judgment only at later stages. Evidence for this second possibility stems from the spillover region (critical word+1), where younger people continued to show longer RTs for sentences with mismatching plural verbs compared to matching singular verbs, while the older group showed the opposite pattern. That is, while the younger group was still preoccupied with the mismatch between the subject and the verb, the older group had already "made up their mind" on the critical word and, having identified the sentence as ungrammatical, quickly moved on to the end-of-trial question. In fact, the verb number by age group interaction was still present in the judgment RTs: the younger group took longer to reject sentences with incorrect plural verbs, while the older group took more time to accept sentences with correct singular verbs. In other words, while the older group decided whether or not a sentence was grammatical upon encountering the incorrect word itself, the younger group used the time between the critical word and the judgment prompt to make this decision. One might speculate that differences in available WM resources might have influenced which of these strategies was more likely to be adopted.

Interestingly, neither self-paced reading latencies nor judgment times were sensitive to agreement attraction. The former showed a main effect of attractor number on the verb, but no interaction between the two; that is, latencies for both matching and mismatching verbs were longer if the verb was preceded by a plural noun. This was unexpected, particularly given that previous comprehension studies have reported attraction effects in German (Häussler, 2009). One explanation for the absence of an judgment RTs: the younger attraction effect in reading times could be that the effects of attractor number (longer reading times for plural attractors) spilled over from the region of the attractor NP to the directly adjacent verb. This may have led to increased reading times for *SPP sentences on the verb (the critical word), yielding similar RTs as for the *SSP condition and thereby masking a possible attraction effect (see Wagers et al., 2009, for a discussion on similar effects).[3]

General discussion

As laid out in the introduction, aging and its effects on language processing have received a growing amount of attention in the past decades. While age differences in lexical processing are relatively robust and well-studied, findings with regard to syntactic processing are mixed. In the present paper, we investigated how aging affects subject-verb agreement processing. In two experiments, we compared a group of younger and a group of older native speakers of German as they processed sentences with attraction errors, in which the finite verb matched in number with an intervening attractor noun rather than with the subject. Experiment 1 established age differences in participants' judgment accuracy. Older people were more prone to incorrectly accepting sentences containing a plural verb following a plural attractor noun compared to younger people. Experiment 2 replicated this finding and additionally found that the online processing of a verb that mismatches the subject in number (compared to a matching verb) is slowed down in older people to a greater extent than in younger speakers. In other words, both experiments found evidence that subject-verb agreement is indeed affected by aging. What is more, Experiment 2 found that the age differences were further modulated by WM capacity within older people. In what follows, we will relate our findings to previous work on the influence of aging and WM on agreement processing and syntactic processing in general.

WM is a limited-capacity memory system whose main function is the temporary maintenance or storage and the manipulation or transformation of information. One of the most influential models of WM stems from Baddeley (Baddeley, 1986, 1992; Baddeley & Hitch, 1974) and consists of three components (central executive, phonological loop, visuospatial sketchpad), all of which have been shown to undergo age-related decline between young adulthood and old age (see Bopp & Verhaeghen, 2005, for a meta-analysis). In the wake of Baddeley's model, some alternative WM models have been proposed more recently which find growing application in the sentence processing literature. The activation-based model of Lewis and Vasishth (2005), for example, assumes that linguistic items are represented as feature bundles in memory. The activation level of the memory traces of these bundles decays over time, which may lead to interference from features of other linguistic items, particularly those that resemble the item in question in some of their features.

A number of studies have investigated the relationship of WM and sentence processing, many of which find that limitations in WM capacities may lead to deficits in the comprehension of complex sentences (King & Just, 1991; Miyake, Carpenter, & Just, 1994; Zurif, Swinney, Prather, Wingfield, & Brownell, 1995). Crucially, however, there seems to be a difference between syntactic and semantic complexity. The extraction of complex semantic information seems to be strongly associated with WM skills (Martin & Feher, 1990; Martin, Shelton, & Yaffee, 1994; Vallar & Baddeley, 1984; Waters, Caplan, & Hildebrandt, 1991; see Caplan & Waters, 1999; and Smith & Geva, 2000; for reviews). Studies investigating the relationship between syntactic complexity and WM, on the other hand, have reported mixed findings. While some studies report preserved processing even when WM abilities were limited (Caplan & Waters, 1999; for a review; Butterworth, Campbell, & Howard, 1986; Butterworth, Shallice, & Watson, 1990; Martin & Feher, 1990; Waters & Caplan, 1996; Waters et al., 1991), others do find limited processing abilities as a function of WM involvement (Friedmann & Gvion, 2002). With respect to agreement processing in particular, Bock and Cutting (1992) measured verbal WM using the speaking-span test (Daneman & Green, 1986) and found that this measure correlated with agreement errors in one of the three experiments they reported. Fayol and colleagues (Fayol et al., 1994; Hupet et al., 1998) found that a secondary task interfered with the processing of subject-verb agreement in French and led to an increase in agreement-attraction errors; see Hartsuiker and Barkhuysen (2006) for similar findings in Dutch.

In light of these findings, our results do not seem surprising. Recall that several accounts have attempted to explain why agreement attraction occurs, with the two most prominent theories being based on "faulty representations" vs. "faulty retrieval." Early theories have attributed attraction errors to feature percolation in which the plural feature of the attractor NP "overwrites" the number feature of the subject NP (i.e., faulty representations). This has been challenged by findings of attraction from nonintervening attractors and "grammaticality asymmetry" (e.g., Wagers et al., 2009), which support an explanation under which attraction is based on simultaneous activation of two NPs and the consequence of faulty retrieval of the NP and its number feature. This latter view has often been tied to the activation-based model of WM (Lewis & Vasishth, 2005; see Tucker et al., 2015, for discussion).

While our study was not designed to decide between these two accounts, it appears that our findings can be most easily explained under the cue-based faulty-retrieval account. First, we found evidence for a "grammaticality asymmetry" – that is, attraction effects emerged for ungrammatical but not for grammatical sentences. As mentioned earlier, percolation models (which attribute attraction to faulty representations of the subject) would predict that an attractor's plural feature percolates upwards to the subject regardless of the following verb. We would then also expect attraction errors to occur in grammatical sentences, that is, incorrect rejections of SPS sentences. Cue-based retrieval accounts, on the other hand, explain the asymmetry by arguing that in cases in which there is a fully-matching NP (Case: nominative, Number: singular), it will outcompete an attractor NP whose features match only partially, so attraction does not occur.

Second, our finding that speakers' susceptibility to attraction errors is modulated by age and by WM favors cue-based retrieval, which is a memory-based account in its very nature. Agreement processing involves keeping the subject and its feature bundle (including its number feature) active in WM during the processing of Participants were asked to subsequent sentence material, retrieving it when encountering the verb, and establishing an agreement dependency between the two if their features match. A sentence's grammatical representation, including the agreement dependency, may need to be reaccessed when readers or listeners are asked to provide an end-of-sentence grammaticality judgment. Older people with lower WM capacity may then experience greater difficulty establishing or recovering grammatical dependencies than high WM people, and as a result may be more prone to interference from an intervening attractor noun. The feature-percolation account, on the other hand, is a purely grammar-based approach to agreement attraction, which would not readily predict age or WM-based effects on speakers' susceptibility to attraction errors.

Interestingly, however, age-related WM differences can only partly account for the observed differences between younger and older speakers. Recall that we found WM differences to modulate older, but not younger participants' performance. In other words, Experiment 2 revealed a cumulative effect of WM capacity and age, rather than age serving as a proxy for WM. Older people with poor WM skills were less likely to notice attraction errors and showed longer reading latencies for mismatching verbs compared with older people with good WM skills. For the younger group, on the other hand, individual WM differences did not seem to matter. Given that our groups differed significantly with regard to their mean WM scores, one possible explanation for this finding is that WM only affects grammatical processing at the lower end of the memory spectrum. However, 72% of the older participants fell into the same range of WM scores as the younger group, suggesting that the difference in the effect size of the WM score was not driven solely by differences in the distribution of scores between the two groups. We suggest that other cognitive capacities that are subject to age-related changes might be involved in the processing of agreement. Regarding the processing of attraction errors in particular, it is conceivable that ignoring an intervening attractor NP draws on inhibitory control skills, which are also known to be subject to age-related decline (Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999). This would be in line with studies finding effects of executive functions and cognitive control for other syntactic phenomena such as ambiguity resolution (see Novick, Trueswell, & Thompson-Schill, 2005, for a review). If the younger group of participants in our study possessed greater inhibitory skills, these might have contributed to more accurate sentence processing by facilitating the blocking of the intervening attractor noun. Further research is necessary to explore the involvement of cognitive control in the processing of attraction errors.

Lastly, we would like to address the issue of differences between online reading latencies and end-of-sentence judgments. As mentioned before, Caplan and Waters (2005) compared the extent to which aging and related differences in WM influence word-by-word listening patterns and end-of-sentence acceptability judgments for sentences containing relative clauses of different degrees of complexity. For their online measure, the authors report no specific effects of age group for two out of three comparisons between different sentence types. For acceptability judgments, however, specific effects of age emerged, with older people showing larger effects of sentence complexity on accuracy rates compared with younger people. Similarly, their participants' reading span scores correlated with judgment accuracy for two of the three comparisons, but only for one comparison for the online listening measure. Caplan and Waters took these differences to mean that if there is a WM system that underlies online syntactic processing, it must be a highly specialized one whose capacity is not measured by standard verbal WM tests. Offline judgments, on the other hand, are affected by more general linguistic processes involving review processes and decision heuristics that are associated with metalinguistic knowledge. In Experiment 2, however, we found age effects in all measures (accuracy and speed of end-of-sentence judgments, reading latencies) as well as modulations of older people's accuracy scores and online reading latencies by WM.

While our participants were not older than the participants in Caplan and Waters (2005), the two experiments differ greatly in terms of the phenomena under study (long-distance movement vs. subject-verb agreement), the modality (auditory vs. visual), and the nature of the task (plausibility judgment vs. grammaticality judgment), all of which may have influenced online reading/listening behavior and Professorship awarded to Harald Clahsen thus the emergence of age and WM effects during online processing.

Conclusion

The current study investigated the extent to which subject-verb agreement computation is affected by aging. Specifically, we compared how younger and older people differ in their susceptibility to agreement attraction errors. Using both self-paced reading and end-of-sentence grammaticality judgments, we measured participants' judgment accuracy, judgment times, and reading latencies. We found that age affected both accuracy and reading latencies. Older people were more likely to accept sentences containing attraction errors than younger people and showed longer reading latencies for verbs that mismatched the subject in number, suggesting that establishing subject-verb agreement dependencies becomes harder with age and more prone to interference. These effects were stronger for older people with low WM scores. Evidently, syntactic processing can be affected by aging, especially when the complexity of the sentence requires the extensive involvement of limited-capacity systems such as WM.

Notes

1. In addition to the experimental items, we further analyzed the grammatically correct filler items to ensure that an effect of verb number was not due to plural forms simply being more difficult to process. We did not find any main effects of age (z = 1.05, *ns*) or verb number (z = 1.36, *ns*), or an interaction between age and verb number (z = 1.32, *ns*) on filler accuracy rates.
2. Note that for easier illustration, working memory was dichotomized for the graphs (Figures 3 and 6) by way of a Median split. All analyses treat working-memory score as a continuous factor, however.
3. We thank an anonymous reviewer for this suggestion.

Acknowledgments

We are grateful to the people who participated in the experiments as well as to three anonymous reviewers for comments on earlier versions of this manuscript. This work was supported by an Alexander-von-Humboldt-Professorship awarded to Harald Clahsen.

Disclosure statement

No potential conflict of interest was reported by the authors.
for accuracy rates and logtransformed

Funding

This work was supported by the Alexander von Humboldt-Stiftung.

References

Baayen, R. H. (2013). *languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics"*. R package version 1.4.1. Retrieved from http://http://CRAN.R-project.org/package=languageR

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559. doi:10.1126/science.1736359

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York, NY: Academic Press.

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. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7. Retrieved from http://arxiv.org/abs/1406.5823

Bennett, I. J., Howard, J. H., & Howard, D. V. (2007). Age-related differences in implicit learning of subtle third-order sequential structure. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *62*, P98–P103. doi:10.1093/geronb/62.2.P98

Bock, K. (2003). Psycholinguistically speaking: Some matters of meaning, marking, and morphing. *Psychology of Learning and Motivation: Advances in Research and Theory*, *44*, 109–144. doi:10.1016/S0079-7421(03)44004-8

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99–127. doi:10.1016/0749-596X(92)90007-K

Bock, J. K., & Eberhard, K. M. (1993). Meaning sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*, 57–99. doi:10.1080/01690969308406949

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*, 45–93. doi:10.1016/0010-0285(91)90003-7

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*, P223–P233. doi:10.1093/geronb/60.5.P223

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in younger and older adults. *Journal of Memory and Language*, *30*, 542–579. doi:10.1016/0749-596X(91)90026-G

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 373–443). New York, NY: Psychology Press.

Butterworth, B., Campbell, R., & Howard, D. (1986). The uses of short-term memory: A case study. *The Quarterly Journal of Experimental Psychology*, *38*, 705–737. doi:10.1080/14640748608401622

Butterworth, B., Shallice, T., & Watson, F. (1990). Short-term retention of sentences without "short-term memory". In G. Vallar & T. Shallice (Eds.), *The neuropsychological impairments of short-term memory* (pp. 187–213). Cambridge: Cambridge University Press.

Caplan, D., & Waters, G. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory*, *13*, 403–413. doi:10.1080/09658210344000459

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*, 77–126. doi:10.1017/S0140525X99001788

Carpenter, P. A., Miyake, A., & Just, M. A. (1994). Working memory constraints in comprehension: Evidence from individual difference, aphasia and aging. In M. A. Gemsbacher (Ed.), *Handbook of psycholinguistics* (pp. 1075–1122). New York, NY: Academic Press.

Churchill, J. D., Stanis, J. J., Press, C., Kushelev, M., & Greenough, W. T. (2003). Is procedural memory relatively spared from age effects? *Neurobiology of Aging*, *24*, 883–892. doi:10.1016/S0197-4580(02)00194-X

Clahsen, H., & Reifegerste, J. (in press). Morphological processing in old-age bilinguals. In M. Libben, T. Gollan, & G. Libben (Eds.), *Bilingualism: A framework for understanding the mental lexicon*. Amsterdam: Benjamins.

Clifton, C., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di lingüística*, *11*, 11–40.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 474–479. doi:10.1037/0278-7393.13.3.474

Cross, E. S., & Burke, D. M. (2004). Do alternative names block young and older adults retrieval of proper names? *Brain and Language*, *89*, 174–181. doi:10.1016/S0093-934X(03)00363-8

Curran, T. (1997). Effects of aging on implicit sequence learning: Accounting for sequence structure and explicit knowledge. *Psychological Research*, *60*, 24–41. doi:10.1007/BF00419678

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, *25*, 1–18. doi:10.1016/0749-596X(86)90018-5

Davis, S. W., Zhuang, J., Wright, P., & Tyler, L. K. (2014). Age-related sensitivity to task-related modulation of language-processing networks. *Neuropsychologia*, *63*, 107–115. doi:10.1016/j.neuropsychologia.2014.08.017

Deutsch, A., & Bentin, S. (1994). Attention mechanisms mediate the syntactic priming effect in auditory word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 595–607. doi:10.1037/0278-7393.20.3.595

Deutsch, A., & Dank, M. (2009). Conflicting cues and competition between notional and grammatical factors in producing number and gender agreement: Evidence from Hebrew. *Journal of Memory and Language*, *60*, 112–143. doi:10.1016/j.jml.2008.07.001

Dobbs, A. R., & Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging*, *4*, 500–503. doi:10.1037/0882-7974.4.4.500

Eberhard, K. M. (1999). The accessibility of conceptual number to the processes of subject-verb agreement in English. *Journal of Memory and Language*, *41*, 560–578. doi:10.1006/jmla.1999.2662

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, *112*, 531–559. doi:10.1037/0033-295X.112.3.531

Fayol, M., Largy, P., & Lemaire, P. (1994). When cognitive overload enhances subject-verb agreement errors: A study in French written language. *Quarterly Journal of Experimental Psychology*, *47*, 437–464. doi:10.1080/14640749408401119

Feeney, J. J., Howard, J. H., & Howard, D. V. (2002). Implicit learning of higher order sequences in middle age. *Psychology and Aging*, *17*, 351–355. doi:10.1037/0882-7974.17.2.351

Feyereisen, P. (1997). A meta-analytic procedure shows an age-related decline in picture naming: Comments on Goulet, Ska, and Kahn (1994). *Journal of Speech, Language, and Hearing Research*, *40*, 1328–1333. doi:10.1044/jslhr.4006.1328

Foos, P. W. (1989). Adult age differences in working memory. *Psychology and Aging*, *4*, 269–275. doi:10.1037/0882-7974.4.3.269

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124. doi:10.3758/BF03195503

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, *23*, 329–374. doi:10.1080/01690960701467993

Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, *17*, 371–404. doi:10.1080/01690960143000254

Frensch, P. A., & Miner, C. S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition*, *22*, 95–110. doi:10.3758/BF03202765

Friedmann, N., & Gvion, A. (2002). Modularity in developmental disorders: Evidence from specific language impairment and peripheral dyslexias. *Behavioral and Brain Sciences*, *25*, 756–757. doi:10.1017/S0140525X02270132

Hartsuiker, R. J., Antón-Méndez, I., & Van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, *45*, 546–572. doi:10.1006/jmla.2000.2787

Hartsuiker, R. J., & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*, 181–204. doi:10.1080/01690960400002117

Hartsuiker, R. J., Kolk, H. H., & Huinck, W. J. (1999). Agrammatic production of subject–verb agreement: The effect of conceptual number. *Brain and Language*, *69*, 119–160. doi:10.1006/brln.1999.2059

Hartsuiker, R. J., Schriefers, H. J., Bock, K., & Kikstra, G. M. (2003). Morphophonological influences on the construction of subject-verb agreement. *Memory & Cognition*, *31*, 1316–1326. doi:10.3758/BF03195814

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). San Diego, CA: Academic.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 653–675). Cambridge, MA: MIT Press.

Haskell, T. R., & MacDonald, M. C. (2003). Conflicting cues and competition in subject-verb agreement. *Journal of Memory and Language*, *48*, 760–778. doi:10.1016/S0749-596X(03)00010-X

Häussler, J. (2009). *The emergence of agreement attraction errors during sentence comprehension* (PhD thesis). Universität Konstanz, Konstanz, Germany.

Heine, M. K., Ober, B. A., & Shenaut, G. K. (1999). Naturally occurring and experimentally induced tip-of-the-tongue experiences in three adult age groups. *Psychology & Aging*, *14*, 445–457. doi:10.1037/0882-7974.14.3.445

Howard, D. V., & Howard, J. H. (1989). Age differences in learning serial patterns: Direct vs. indirect measures. *Psychology and Aging*, *4*, 357–364. doi:10.1037/0882-7974.4.3.357

Howard, D. V., & Howard, J. H. (1992). Adult age differences in the rate of learning serial patterns: Evidence from direct and indirect tests. *Psychology and Aging*, *7*, 232–241. doi:10.1037/0882-7974.7.2.232

Howard, D. V., Howard Jr., J. H., Japikse, K., DiYanni, C., Thompson, A., & Somberg, R. (2004). Implicit sequence learning: Effects of level of structure, adult age, and extended practice. *Psychology and Aging*, *19*, 79–92. doi:10.1037/0882-7974.19.1.79

Howard Jr., J. H., Howard, D. V., Dennis, N. A., Yankovich, H., & Vaidya, C. J. (2004). Implicit spatial contextual learning in healthy aging. *Neuropsychology*, *18*, 124–134. doi:10.1037/0894-4105.18.1.124

Humphreys, K. R., & Bock, K. (2005). Notional number agreement in English. *Psychonomic Bulletin & Review*, *12*, 689–695. doi:10.3758/BF03196759

Hupet, M., Fayol, M., & Schelstraete, M. A. (1998). Effects of semantic variables on the subject-verb agreement processes in writing. *British Journal of Psychology*, *89*, 59–75. doi:10.1111/j.2044-8295.1998.tb02673.x

Jespersen, O. (1924). *The philosophy of grammar*. London: Allen & Unwin.

Just, M., Carpenter, P., & Woolley, J. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*, 228–238. doi:10.1037/0096-3445.111.2.228

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149. doi:10.1037/0033-295X.99.1.122

Kaan, E. (2002). Investigating the effects of distance and number interference in processing subject-verb dependencies: An ERP study. *Journal of Psycholinguistic Research*, *31*, 165–193. doi:10.1023/A:1014978917769

Kavé, G., & Levy, Y. (2005). The processing of morphology in old age: Evidence from Hebrew. *Journal of Speech, Language, and Hearing Research*, *48*, 1442–1451. doi:10.1044/1092-4388(2005/100)

Kemper, S. (1986). Imitation of complex syntactic constructions by elderly adults. *Applied Psycholinguistics*, *7*, 277–287. doi:10.1017/S0142716400007578

Kemper, S., Herman, R. E., & Lian, C. H. (2003). The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech of noise. *Psychology and Aging*, *18*, 181–192. doi:10.1037/0882-7974.18.2.181

Kemtes, K. A., & Kemper, S. (1997). Younger and older adults' on-line processing of syntactically ambiguous sentences. *Psychology and Aging*, *12*, 362–371. doi:10.1037/0882-7974.12.2.362

Kimball, J., & Aissen, J. (1971). I think, you think, he think. *Linguistic Inquiry*, *2*, 241–246.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580–602. doi:10.1016/0749-596X(91)90027-H

Lago, S., Shalom, D., Sigman, M., Lau, E., & Phillips, C. (2014). Agreement processes in Spanish comprehension. *Journal of Memory and Language*, *82*, 133–149. doi:10.1016/j.jml.2015.02.002

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging*, *8*, 34–43. doi:10.1037/0882-7974.8.1.34

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*, 375–419. doi:10.1207/s15516709cog0000_25

Light, L. L., & Anderson, P. A. (1985). Working-memory capacity, age, and memory for discourse. *Journal of Gerontology*, *40*, 737–747. doi:10.1093/geronj/40.6.737

Lorimor, H., Bock, K., Zalkind, E., Sheyman, A., & Beard, R. (2008). Agreement and attraction in Russian. *Language and Cognitive Processes*, *23*, 769–799. doi:10.1080/01690960701774182

Mallinson, G., & Blake, B. J. (1981). *Language typology. Cross-linguistic studies in syntax*. Amsterdam: North-Holland.

Martin, E., & Roberts, K. H. (1967). Sentence length and sentence retention in the free-learning situation. *Psychonomic Science*, *8*, 535–536. doi:10.3758/BF03331738

Martin, R. C., & Feher, E. (1990). The consequences of reduced memory span for the comprehension of semantic versus syntactic information. *Brain and Language*, *38*, 1–20. doi:10.1016/0093-934X(90)90099-3

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, *33*, 83–111. doi:10.1006/jmla.1994.1005

Meinzer, M., Flaisch, T., Seeds, L., Harnish, S., Antonenko, D., Witte, V., … Crosson, B. (2012a). Same modulation but different starting points: Performance modulates age differences in inferior frontal cortex activity during word-retrieval. *PLoS One*, *7*, e33631. doi:10.1371/journal.pone.0033631

Meinzer, M., Seeds, L., Flaisch, T., Harnish, S., Cohen, M. L., McGregor, K., … Crosson, B. (2012b). Impact of changed positive and negative task-related brain activity on word-retrieval in aging. *Neurobiology of Aging*, *33*, 656–669. doi:10.1016/j.neurobiolaging.2010.06.020

Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology*, *11*, 671–717. doi:10.1080/02643299408251989

Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject-verb agreement processes in comprehension. *Journal of Memory and Language*, *36*, 569–587. doi:10.1006/jmla.1996.2497

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *5*, 263–281. doi:10.3758/CABN.5.3.263

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N., Smith, A. D., & Smith, P. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*, 299–320. doi:10.1037/0882-7974.17.2.299

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427–456. doi:10.1006/jmla.1999.2653

Poeppel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language*, *69*, 1–33. doi:10.2307/416414

Prull, M. W., Gabrieli, J. D. E., & Bunge, S. A. (2000). Age-related changes in memory: A cognitive neuroscience perspective. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 91–153). Mahwah, NJ: Lawrence Erlbaum.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A grammar of contemporary English*. London: Longman.

R Development Core Team. (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Rabinowitz, J. C. (1984). Aging and recognition failure. *Journal of Gerontology*, *39*, 65–71. doi:10.1093/geronj/39.1.65

Rastle, K. G., & Burke, D. M. (1996). Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. *Journal of Memory and Language*, *35*, 586–605. doi:10.1006/jmla.1996.0031

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428. doi:10.1037/0033-295X.103.3.403

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology*, *54*, 35–54. doi:10.1016/S0301-0511(00)00052-1

Schmitter-Edgecombe, M., Vesneski, M., & Jones, D. W. R. (2000). Aging and word-finding: A comparison of spontaneous and constrained naming tests. *Archives of Clinical Neuropsychology*, *15*, 479–493. doi:10.1016/S0887-6177(99)00039-6

Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P. P., & Tyler, L. K. (2007). On the tip-of-the-tongue: Neural correlates of increased word-finding failures in normal aging. *Journal of Cognitive Neuroscience*, *19*, 2060–2070. doi:10.1162/jocn.2007.19.12.2060

Shafto, M. A., & Tyler, L. K. (2014). Language in the aging brain: The network dynamics of cognitive decline and preservation. *Science*, *346*, 583–587. doi:10.1126/science.1254404

Smith, A. D. (1977). Adult age differences in cued recall. *Developmental Psychology*, *13*, 326–331. doi:10.1037/0012-1649.13.4.326

Smith, E. E., & Geva, A. (2000). Verbal working memory and its connections to language processing. In Y. Grodzinsky, L. P. Shapiro, & D. Swinney (Eds.), *Language and the brain* (pp. 123–141). San Diego: Academic Press.

Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing number in the real-time processing of agreement: Self-paced reading evidence from Arabic. *Frontiers in Psychology*, *6*, 347. doi:10.3389/fpsyg.2015.00347

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: The modulation of the frontotemporal language system in the context of age-related atrophy. *Cerebral Cortex*, *20*(2), 352–364. doi:10.1093/cercor/bhp105

Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior*, *23*, 151–161. doi:10.1016/S0022-5371(84)90104-X

Verhaeghen, P., Marcoen, A., & Goossens, L. (1993). Facts and fiction about memory aging: A quantitative integration of research findings. *Journal of Gerontology*, *48*, P157–P171. doi:10.1093/geronj/48.4.P157

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age–cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin*, *122*, 231–249. doi:10.1037/0033-2909.122.3.231

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject–verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, *61*, 261–298. doi:10.1016/S0010-0277(96)00713-5

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject–verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, *34*, 186–215. doi:10.1006/jmla.1995.1009

Vigliocco, G., Hartsuiker, R. J., Jarema, G., & Kolk, H. H. J. (1996). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes*, *11*, 407–442. doi:10.1080/016909696387169

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition*, *68*, B13–B29. doi:10.1016/S0010-0277(98)00041-9

Wagers, M., Lau, E., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*, 206–237. doi:10.1016/j.jml.2009.04.002

Waters, G., Caplan, D., & Hildebrandt, N. (1991). On the structure of verbal short-term memory and its functional role in sentence comprehension: Evidence from neuropsychology. *Cognitive Neuropsychology*, *8*, 81–126. doi:10.1080/02643299108253368

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology*, *49*, 51–74. doi:10.1080/713755607

Wearing, A. J. (1969). Recognition memory for sentences of varying length. *Psychonomic Science*, *15*, 221–222. doi:10.3758/BF03336293

Zurif, E., Swinney, D., Prather, P., Wingfield, A., & Brownell, H. (1995). The allocation of memory resources during sentence comprehension: Evidence from the elderly. *Journal of Psycholinguistic Research*, *24*, 165–182. doi:10.1007/BF02145354