Psicothema

Psicothema 2020, Vol. 32, No. 3, 307-313
doi: 10.7334/psicothema2020.90

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

Phonological false recognition, recollection, and familiarity in healthy aging and Alzheimer's disease

Alfonso Pitarque[1], Encarnación Satorres[1], Joaquín Escudero[2], Salvador Algarabel[1], and Juan C. Meléndez[1]
[1] University of Valencia, and [2] General Hospital (Valencia)

Abstract

Background: Thirty healthy older people, 20 Alzheimer's disease patients (matched on age and education level) and 33 young people, participated in an experiment to implicitly induce phonological false memories, allowing us to obtain estimates of their recollection, familiarity, and false recognition. **Method:** In the study task, words were selected which used half of the letters in the alphabet. In the recognition test, there were three types of non-studied new words: critical lures using letters from the same half of the alphabet as the study task words; distractors formed using the unused half of the alphabet, and distractors formed using all the letters in the alphabet. **Results:** Results showed that: (a) in all the samples, critical lures produced more false recognitions than distractors composed of all the letters in the alphabet or distractors composed of the letters not used in the study, showing a significant phonological false recognition effect; (b) both recollection and familiarity declined with age and dementia; (c) phonological false recognition increased with age and Alzheimer's disease. **Conclusions:** These results seem to support the idea that estimates of recollection, familiarity, and phonological false recognition can be used as early markers of cognitive impairment.

Keywords: Phonological false recognition; recollection; familiarity; aging; Alzheimer's disease.

Resumen

Falso reconocimiento fonológico, recolección y familiaridad en el envejecimiento y la enfermedad de Alzheimer. **Antecedentes:** treinta personas mayores sanas, 20 pacientes con enfermedad de Alzheimer (igualados en edad y nivel educativo) y 33 jóvenes participaron en un experimento para inducirles implícitamente falsas memorias fonológicas, permiténdonos obtener sus estimaciones de recolección, familiaridad y falso reconocimiento. **Método:** en la tarea de estudio las palabras estaban formadas por una mitad de las letras del alfabeto. En el test de reconocimiento había tres tipos de palabras nuevas no estudiadas: palabras críticas formadas por las mismas letras de la tarea de estudio, distractores formados por la otra mitad de letras no utilizadas en la tarea de estudio y distractores formados por todas las letras del alfabeto. **Resultados:** los resultados mostraron que: (a) en las tres muestras las palabras críticas producían más falsos reconocimientos que en ambos tipos de distractores, mostrando un claro efecto significativo de falso reconocimiento fonológico; (b) tanto la recolección como la familiaridad disminuían durante el envejecimiento y la enfermedad de Alzheimer; (c) el falso reconocimiento fonológico aumentaba durante el envejecimiento y la enfermedad de Alzheimer. **Conclusiones:** nuestros resultados apoyan la idea de que las estimaciones de recolección, familiaridad y falso reconocimiento fonológico pueden ser utililizadas como marcadores tempranos de deterioro cognitivo.

Palabras clave: falso reconocimiento fonológico; recolección; familiaridad; envejecimiento; enfermedad de Alzheimer.

It has been well established that healthy aging, Mild Cognitive Impairment (MCI), and Alzheimer's disease (AD) are associated with a decline in episodic memory (Koen & Yonelinas, 2014) and an increase in false memories (Devitt & Schacter, 2016). However, there is still an intense debate about how two forms of episodic memory (recollection and familiarity) are affected by healthy and pathological aging. Our recall or recognition of a past experience can be based on either a conscious *recollection* of contextual details from that experience or on an automatic estimation of the strength of this memory trace in the absence of contextual details (*familiarity*). Neuroimaging studies suggest that these two processes seem to rest on different neuroanatomical bases within the medial temporal lobe: recollection seems to be related to the functioning of the hippocampus (Schoemaker et al., 2017), whereas familiarity is associated with the perirhinal and entorhinal cortices (Brandt et al., 2016). Recently Scalici et al. (2017), in a review of studies using functional magnetic resonance imaging, also extended these dissociations to different areas of the prefrontal cortex, whereas King et al., (2018) located these dissociations within the striatal region. For these reasons, the models explaining these dissociations are generically referred to as *dual-process* theories (Koen & Yonelinas, 2014; Schoemaker et al., 2014), receiving abundant experimental support (Koen & Yonelinas, 2014; Schoemaker et al., 2014), even in animal cognition (Basile & Hampton, 2013).

Received: March 11, 2020 • Accepted: May 19, 2020
Corresponding author: Alfonso Pitarque
Faculty of Psychology
University of Valencia
46010 Valencia (Spain)
e-mail: pitarque@uv.es

Two recent literature reviews have shown that there seems to be unanimity in accepting that recollection tends to decline with age (Koen & Yonelinas, 2014) and cognitive impairment (Schoemaker et al., 2014), to such an extent that recollection deficits are usually interpreted as the main prodromal markers of some of the most prevalent cognitive pathologies in old age (Schoemaker et al., 2017). However, the experimental results related to familiarity are not as conclusive, and they seem to differ depending on the experimental paradigm used. Thus, in healthy aging, familiarity was not impaired in studies using the process-dissociation (PD) procedure or the receiver operating characteristics (ROC) procedure, but it was impaired in studies that used the remember-know (RK) procedure (Koen & Yonelinas, 2014, Pitarque et al., 2015). However, Koen & Yonelinas (2016) did not actually find any impairment in familiarity with the RK procedure. With regard to familiarity in participants with cognitive impairment, results are even less consistent. For example, in their review, Schoemaker et al. (2014) found six studies that demonstrated a decline in familiarity-based recognition in MCI individuals, whereas five studies pointed to its preservation. In this regard, some authors suggest using these deficits in familiarity as another early cognitive prodromal marker of AD (Wolk et al., 2013). Thus, the results of the review by Schoemaker et al., (2014) seem to demonstrate that familiarity and recollection are differentially affected by advancing AD neuropathology. Whereas recollection is broadly affected throughout all the stages of the disease process, familiarity deficits seem to be present only in more advanced stages of cognitive impairment, which would be consistent with dual-process theories. Therefore, more research is necessary to analyze the roles that familiarity and recollection play in cognitive impairment, and this is the first objective of our study.

The second objective of our study is to analyze the role of healthy and neuropathological aging in false recognition. It has been widely accepted that false memories or false recognitions (as well as general semantic and phonological errors, omissions, perseverations, illusions, etc.) increase with age and cognitive impairment (Devitt & Schacter, 2016; McCabe et al., 2009). Thus, false recognition has also been used as another prodromal marker of neuropathological aging (Hildebrant et al., 2009). Traditionally, false memories have been studied using the Deese-Roediger-McDermott (DRM) procedure (Roediger & McDermott, 1995). In this paradigm, the stimuli studied are semantically related to each other (for example, "tiger", "puma", "cat"), which can cause the false recognition of critical stimuli not studied but semantically related to the study list (for example, "panther"). The increase in false memories with healthy or pathological aging has traditionally been explained in the literature mainly by two theoretical models. On the one hand, the *fuzzy-trace* theory (Reyna & Brainerd, 1995) emphasizes the fact that old people, due to their limited capacity to recollect item-specific information (representations containing perceptual and contextual details), tend to trust their retrieval judgments in their *gist* memory (or semantic-based information underlying the stimuli studied), producing an increase in their false recognition. That is, from this perspective, true memory of presented items relies on the presence of both item-specific information and gist traces. In contrast, false recognition is attributed to the persistence of the gist representation during the recognition phase, as well as inability to use item-specific information of truly presented words to suppress the acceptance of critical distractors (see e.g. Rodríguez-Ferreiro et al., 2019, with results supporting this theory). On the other hand, the dual *activation-monitoring* theory (Gallo, 2010; Roediger et al., 2001), establishes that, during the study task, both studied items and items semantically related to them are activated (because activation spreads from one to the other), making them more accessible for later retrieval. At the time of recognition, the subject carries out a conscious monitoring process to distinguish between studied and non-studied items. Given that critical items (that is, the new items related to the studied items) can be highly activated, source-monitoring errors can occur (Johnson et al., 1993), leading to false memories or false recognitions. However, because young adults have a well-preserved capacity to recollect item-specific contextual information, they can use conscious monitoring strategies such as *recall-to-reject* to reduce their false alarm rates (e.g. "I know I did not study panther because I remember that I studied tiger"; Brainerd et al., 2003; see also Basile & Hampton, 2013, regarding its use in animal cognition). In other words, whereas activation enhances false memories, monitoring reduces them (Gallo, 2010). In this regard, Schoemaker et al. (2017) recently showed that the increased use of familiarity-based recognition in older adults is positively associated with an increase in false alarms, whereas recollection does not seem to be associated with false recognition. However, in their meta-analytic review of the RK experimental paradigm, McCabe et al. (2009) showed that age-related deficits are also observed in an increase in false alarms on recollection-type judgements, which would indicate that these deficits are also due to an incorrect recollection of episodic traces associated with other items studied earlier, rather than to an incorrect use of familiarity. Due to these inconsistent results, it is also necessary to analyze the role of healthy and neuropathological aging in false recognition, which is the second objective of our study.

Various studies have shown that it is also possible to elicit false memories of critical words (e.g., "chair") after studying words related to them phonologically rather than semantically (e.g., "cheer", "hair"; Finley et al., 2017; Watson et al., 2003). These phonological false memories have also been explained by the activation-monitoring theory. This model assumes that studying a word produces the activation of the phonemes that make it up, and this activation is propagated bottom-up toward other words that share these phonemes, increasing their activation and making them more accessible to later recall or recognition. Additional support for the activation-monitoring theory stems from the fact that combining phonological and semantic associates on the study lists produces over-additive effects on false memory (Finley et al., 2017; Watson et al., 2003). However, the fuzzy-trace theory cannot explain phonological false memories because as they are entirely based on a gist trace (a semantic content that does not exist in the phonological false recognition paradigm because the study words are not semantically related) there is no reason to expect phonological associates to increase false recall or false recognition (Finley et al., 2017).

Phonological false memories seem to increase with healthy aging and AD in a similar way to semantic false memories (Sommers & Huff, 2003; Watson et al., 2001). This phenomenon has been interpreted as a poorer use of monitoring strategies by older adults and AD patients, which reduces their ability to inhibit activated lexical competitors (Finley et al., 2017; Sommers & Huff, 2003). However, Budson et al., (2003) found that older adults showed higher levels of phonological false recognition than younger adults, whereas AD patients exhibited lower levels of

false recognition than older adults. This result was explained by the episodic memory and attentional control system deficits shown by AD patients. Therefore, it also seems theoretically relevant to analyze the role played by aging and dementia in phonological false memories.

For the aforementioned reasons, we propose an experiment to elicit phonological false memories based on a perceptual manipulation of the stimuli, which is implicit for the participants, in order to increase the activation of critical words (see Pitarque et al., 2019). This new procedure, adapted from the one proposed by Parkin et al., (2001), mainly consists of presenting study words formed from half of the letters in the alphabet. On the subsequent recognition test, the new words can be formed either from the same letters as the studied words (or *critical* lures because they are phonologically related to the studied words), distractors formed from the other half of the letters in the alphabet, or distractors formed from the entire alphabet. On the recognition test, after each "old" response, participants have to emit a second introspective judgment about whether their positive recognition was based on the recollection of episodic details associated with the item (R judgment) or on a mere sensation of familiarity with the item in the absence of episodic details (K judgment). Based on the individual rates of hits and false alarms for R and K judgments, we estimate the individual recollection and familiarity rates (Koen & Yonelinas, 2016) and the relative false recognition rates (Budson et al., 2003; Watson et al., 2001). Thus, with this new paradigm, we compare a sample of young people, a sample of healthy older people, and a sample of AD patients (matched in age and education level to the sample of older people) to analyze the effects of healthy aging and Alzheimer's disease on recollection, familiarity, and phonological false recognition.

Method

Participants

The sample of young people was composed of 33 Psychology degree students from the University of Valencia (26 women, 7 men), ranging from 18 to 30 years old (M = 20.73, SD = 3.19). The sample of older people consisted of 30 older participants (20 women, 10 men), ranging from 65 to 83 years old (M = 69.47, SD = 4.10), who belonged to various leisure centers for older people in the city of Valencia, Spain. These participants reported being in good physical and mental health, with no known memory impairments. In this regard, the mean for the older people on the Mini-Mental State Examination (MMSE; Folstein et al., 1975) was 28.53 (SD = 1.28, range 26-30), revealing no memory impairment. The sample of AD patients was composed of 20 people (7 women, 13 men), ranging from 55 to 76 years old (M = 66.78, SD = 6.81), who were patients from the Department of Neurology at the General Hospital of Valencia with a clinical diagnosis of probable AD. The AD patients' mean score on the MMSE was 23.50 (SD = 1.85, range 21-27), significantly below the mean of the healthy older people ($t(48) = 11.39$, $p < .0001$). In the studies reviewed by Schoemaker et al. (2014; and also Koen & Yonelinas, 2014), the MMSE mean ranged from 18.6 to 25.6 for the AD groups and from 25.5 to 28.5 for the MCI groups. Thus, from this point of view, our patients could be considered clinical AD patients. In addition, all our AD patients had PET images with positive amyloid compatible with Alzheimer's disease (Dubois et al., 2016). The samples of older people and AD patients were matched on age ($t(48) = 1.83$, $p = .11$) and education level (using an ordinal four-point scale, ranging from 1 = *basic education* to 4 = *university education*; Mann-Whitney's $z = 0.18$, $p = .86$). AD patients were diagnosed based on the criteria described by Dubois et al. (2016). All the AD patients underwent extensive evaluation by a neurologist and a neuropsychologist, including physical and neurological exams, history from both patient and informant, psychometric testing and amyloid PET scanning. Participants were excluded if they had a history of clinical stroke, traumatic brain injury, alcohol or drug abuse/dependence, prior electroconvulsive therapy, and any significant disease or medical/psychiatric condition that might impact neuropsychological performance. The study was approved by the Institutional Ethics Committee of the University of Valencia. All the participants voluntarily gave their written consent to participate.

Instruments

We used three lists of 46 words each (see Table 1): words of List A were formed entirely with the following letters from the Spanish alphabet: a, e, u, b, d, g, j, n, r, z; words of List B were formed entirely with the following letters i, o, c, f, h, l, m, p, s, t, v, y; words of List C were formed with letters from the entire alphabet, with the only criterion being that each word had to contain at least one letter from List A and at least one letter from List B (Pitarque et al., 2019). Lists A, B, and C were balanced on mean frequency per two million (Alameda & Cuetos, 1995), 82.41 (SD = 158.12), 83.48 (SD = 124.71), and 82.39 (SD = 156.32), respectively ($F(2,135) < 1$), and length, 5.11 (SD = 1.12), 4.74 (SD = 1.31), and 5.07 letters (SD = 1.10), respectively ($F(2,135) = 1.35$, $p = .26$, $\eta^2_p = 0.02$). Words from List A and List B (counterbalanced across participants) were used as study words and old items, critical words and distractors on the recognition task. Words from List C were only used as distractors on the recognition task.

Procedure

The participants, seated in front a computer, performed individually a study and recognition task that took about 15 minutes. On the study task, the participants studied 34 words formed from half of the letters in the alphabet (taken from either list A or List B, counterbalanced across participants). Study words were presented one by one for two seconds each in the center of a white computer screen (using a black Calibri font, 48 points) with an inter-stimuli period of one second. On the subsequent self-paced recognition test, the participants had to recognize 66 words one by one (pressing each time to respond one of two keys labeled on the computer keyboard as "old" or "new"): 30 from those studied previously (e.g., from list A; the first and last two words on each study list were not tested on the later recognition task to avoid effects of primacy and recency, respectively), 12 critical lures formed from the same set of letters as the studied words (e.g., from list A), 12 distractors formed from the remaining letters not used in the study task (e.g., from list B), and 12 distractors formed from the entire alphabet (from list C). For each participant, both the words on the study list and the new words on the recognition task were selected randomly from lists A, B, and C, and presented by the E-prime software for experimental control (Schneider et al., 2002).

Table 1
Lists of words used in the experiment (and English translation)

LIST A	LIST B	LIST C
arruga (wrinkle)	voto (vote)	comarca (region)
barrera (barrier)	móvil (mobile)	retina (retina)
árabe (Arab)	plomo (lead)	pastor (shepherd)
naranja (orange)	tos (cough)	sopa (soup)
aguja (needle)	timo (scam)	locutor (announcer)
danza (dance)	óptico (optician)	trapo (cloth)
granja (farm)	socio (partner)	colilla (cigarette end)
baranda (railing)	olivo (olive tree)	tubo (pipe)
agenda (diary)	tic (tic)	jamón (ham)
rana (frog)	cosmos (cosmos)	atasco (jam)
zanja (ditch)	mili (military service)	aseo (toilet)
ajedrez (chess)	filo (edge)	volante (wheel)
agua (water)	sol (sun)	año (year)
guerra (war)	tío (uncle)	amor (love)
duda (doubt)	motivo (reason)	idea (idea)
edad (age)	chico (boy)	sujeto (subject)
bar (pub)	ocho (eight)	hogar (home)
red (net)	piso (flat)	borde (edge)
arena (sand)	civil (civil)	lujo (luxury)
barra (bar)	mito (myth)	ropa (clothes)
deber (duty)	oficio (job)	baño (bathroom)
azar (random)	otoño (fall)	montaña (mountain)
gana (wish)	foto (photo)	alivio (relief)
guarda (guard)	filósofo (philosopher)	piano (piano)
bandera (flag)	pico (peak)	bosque (forest)
barba (beard)	polo (pole)	baile (dance)
eje (axis)	oso (bear)	beso (kiss)
raza (race)	hocico (snout)	calcio (calcium)
juez (judge)	positivo (positive)	palo (stick)
banda (band)	ciclo (cycle)	remo (rowing)
nube (cloud)	colmo (last straw)	tallo (stem)
rueda (wheel)	vicio (addiction)	corteza (bark)
nuera (daughter in law)	tópico (cliche)	chaleco (vest)
rareza (rarity)	hipo (hiccup)	monje (monk)
ranura (groove)	lomo (loin)	ático (attic)
juerga (party)	físico (physicist)	tinta (ink)
ajuar (trousseau)	ocio (leisure)	imán (magnet)
abeja (bee)	lío (mess)	reto (challenge)
duna (dune)	loco (mad)	velo (veil)
aduana (customs)	tipo (type)	gorra (cap)
andén (platform)	pitillo (cigarette)	tiza (chalk)
adrede (intentionally)	moño (bun)	cesto (basket)
ganga (bargain)	mimo (caress)	corán (koran)
garra (claw)	misil (missile)	talón (heel)
brebaje (potion)	colmillo (tusk)	bombo (drum)
daga (dagger)	olmo (elm)	zumo (juice)

On the recognition test, after each positive recognition ("old" response), the participants had to emit a second introspective self-paced judgment about whether his/her "old" response was based on the recollection of episodic details associated with the item (R judgment) or on a mere sensation of familiarity with the item in the absence of episodic details (K judgment), pressing one of two keys labeled on the computer keyboard as "remember" or "know". If necessary, the experimenter helped elderly and patients to press the answer keys. Prior to performing the first recognition task, and following strict remember-know instructions (Koen & Yonelinas, 2014, 2016), the difference between "remembering" and "knowing" was explained to participants, emphasizing that an R response should only be given if they could communicate a retrieved detail to the experimenter if asked, whereas a K response should be given if they believed the word was previously studied but could not retrieve any specific details about it. A short practice task was performed to make sure all the subjects understood the instructions.

Finally, a debriefing questionnaire asked the participants if they were aware of any relationship between the words in the study task, and they were removed from the study if they answered affirmatively (3 young people were removed from the final sample because they were aware of some kind of phonological relationship between the words).

Data analysis

First, we carried out a mixed ANOVA on false alarms with 3 groups X 3 types of new words to analyze whether our paradigm was able to elicit a phonological false recognition effect. Second, we calculated the recollection and familarity estimates for each participant. The recollection estimates were derived by subtracting the proportion of false alarms on R judgments (FAR) from the proportion of hits on R judgments (HR), whereas familiarity estimates were derived with the formula HK/(1-HR) – (FAK/(1-FAR), where HK is the proportion of hits on K judgments, and FAK is the proportion of false alarms on K judgments (Koen & Yonelinas, 2016). Third, we calculated the relative false recognition estimates (Budson et al., 2003; Watson et al., 2001) by dividing, for each participant, the proportion of false alarms on critical lures by the proportion of hits (Table 2), as a way to control the response bias. Finally, these recollection, familarity and relative false recognition estimates were analyzed by means of three one-way between-subjects analysis of variance (ANOVA) comparing the three groups. Subsequently, Bonferroni post-hoc tests were used to examine differences across groups. A value of $p \leq .05$ was considered statistically significant.

Results

The overall results of our experiment are shown in Table 2. First, to analyze whether our paradigm has enough sensitivity to elicit the phonological false recognition effect, we carried out a mixed ANOVA on false alarms with 3 groups X 3 types

Table 2
Means (and SE) of hits, false alarms (FA), and estimations of recollection, familiarity, and relative false recognition, as a function of groups

	Young	Older people	AD patients
Hits	0.84 (0.02)	0.74 (0.02)	0.75 (0.03)
FA (critical lures)	0.28 (0.03)	0.33 (0.03)	0.46 (0.04)
FA (different letters from the study list)	0.05 (0.02)	0.05 (0.02)	0.23 (0.03)
FA (all the letters of the alphabet)	0.07 (0.02)	0.12 (0.02)	0.26 (0.03)
Recollection	0.64 (0.03)	0.53 (0.03)	0.46 (0.04)
Familiarity	0.41 (0.04)	0.27 (0.04)	0.20 (0.05)
Relative false recognition	0.33 (0.04)	0.46 (0.04)	0.59 (0.05)

of new words (Table 2) that showed that the main effects of the two variables were significant ($F(2, 80) = 11.09$, $p < .001$, $\eta^2_p = 0.22$; $F(2, 160) = 108.92$, $p < .001$, $\eta^2_p = 0.58$, respectively). Post-hoc Bonferroni t-tests comparing the means of the three groups showed that AD patients ($M = .32$) committed more false alarms than young people ($M = .14$; $p < .001$) and healthy older people ($M = .17$; $p < .001$), with no differences between the latter two samples ($p = .33$). On the other hand, post-hoc Bonferroni t-tests comparing the means of the 3 types of new words showed that new words with the same letters as the study list (or critical words; $M = .36$) gave rise to more false alarms than new words with different letters from the study list ($M = .11$; $p < .001$) or new words formed from all the letters in the alphabet ($M = .15$; $p < .001$), being also significant the difference between these two latter conditions ($p = .002$). Finally, the interaction between types of new words and groups was not significant ($F(4, 160) < 1$), indicating that, in the three samples, the critical words elicit more false alarms than the rest of the conditions, showing a significant phonological false recognition effect (as in Pitarque et al., 2019).

The estimations of recollection and familiarity are shown in Table 2. The recollection estimates were analyzed by means of a one-way between-subjects ANOVA comparing the means of the 3 groups, which showed that the effect of this variable was significant ($F(2, 80) = 4.55$, $p = .013$, $\eta^2_p = 0.10$). Post-hoc Bonferroni t-tests comparing the means of the three groups showed that young people ($M = .64$) used recollection more than older people ($M = .53$; $p = .014$), and AD patients ($M = .46$; $p = .005$), being also significant the difference between these two latter conditions ($p = .032$), as commonly found in the literature (Koen & Yonelinas, 2014; Schoemaker et al., 2014). These results show, then, that recollection declines significantly with both healthy and neuropathological aging, thus confirming the idea that recollection deficits can be used as prodromal markers of AD (Schoemaker et al., 2014, 2017).

With regard to the familiarity estimates (Table 2), a one-way between-subjects ANOVA showed that the effect of the group variable was significant ($F(2, 80) = 7.25$, $p = .001$, $\eta^2_p = 0.15$). Post-hoc Bonferroni t-tests comparing the means of the three groups showed that young people ($M = .41$) used familiarity more than older people ($M = .27$; $p = .015$) and AD patients ($M = .18$; $p < .001$), being also significant the difference between these two latter conditions ($p = .043$). These results show, then, that familiarity also declines significantly with both healthy and neuropathological aging, confirming the idea that familiarity deficits can also be used as prodromal markers of AD, as Wolk et al. (2013) suggest.

With regard to the false recognition estimates (Table 2), the one-way ANOVA comparing the means of the 3 groups was significant ($F(2, 80) = 7.16$, $p = .001$, $\eta^2_p = 0.15$). Post-hoc Bonferroni t-tests comparing the means of the three groups showed that AD patients ($M = .59$) committed more phonological false memories than older people ($M = .46$; $p = .044$) and young people ($M = .33$; $p < .001$), being also significant the difference between these two latter conditions ($p = .034$), as commonly found in the literature (Summer & Huff, 2003; Watson et al., 2001), suggesting that false recognition can be considered another prodromal marker of neuropathological aging (Hildebrant et al., 2009).

Discussion

First, our results showed that in all the samples critical lures produced more false recognitions than distractors composed of all the letters in the alphabet or distractors composed of the letters not used in the study, showing a significant phonological false recognition effect. These results can be explained by the activation-monitoring theory (Gallo, 2010; Roediger et al., 2001) which assumes that studying a word produces the activation of the phonemes that make it up, and this activation is propagated bottom-up toward other words that share these phonemes, increasing their activation and making them more accessible to later recognition. Because our critical words share the same letters as the studied words formed with half the letters of the alphabet, they are more activated than the distractors formed by all the letters of the alphabet or distractors formed by the other half of non-studied letters, thus giving rise to more false alarms. Because the AD patients and, to a lesser extent, older people have a limited ability to inhibit activated lexical competitors (Finley et al., 2017; Sommers & Huff, 2003), they commit more false alarms than young people. However, our results cannot be explained by the fuzzy-trace theory because in our experiment the study words were not semantically related, so the participants could not build a gist trace to explain their false recognitions (Finley et al., 2017).

Second, our results also showed that both recollection and familiarity decline with age and dementia. Our results for familiarity seem to refute the dual-process theories (Schoemaker et al., 2014), which argue that whereas recollection is broadly affected throughout all the stages of the disease process, familiarity deficits seem to be present only in more advanced stages of cognitive impairment. However, our results show that recollection and familiarity decrease similarly in healthy aging and dementias. Other authors also found similar results to ours using the PD paradigm in associative recognition (Wolk et al., 2011), the ROC paradigm (Ally et al., 2009), or the RK paradigm (Pitarque et al., 2015). A possible explanation for our results for familiarity could be related to the characteristics of our experimental task, which could produce greater implicit activation of the studied words, given that in our study task, each word is activated by 34 other study words related phonologically to it (a significantly longer list than the study lists commonly used in the DRM paradigm, where each critical word is associated with a maximum of 12 study words; Budson et al., 2003; Watson et al., 2001). Another fact that supports this idea is that in our results, the effect size of familiarity ($\eta^2_p = 0.15$) is larger than the effect size of recollection ($\eta^2_p = 0.10$), unlike what usually occurs on conventional recognition tasks (see e.g. Koen & Yonelinas, 2014). This result would indicate that, in our paradigm, true recognition due to familiarity has greater relevance than true recognition through recollection. That is, our paradigm seems to be more sensitive than others to detecting changes in familiarity due to healthy and pathological aging. Finally, some authors (e.g. Schoemaker et al., 2017) propose the idea that healthy older people and MCI and AD patients increase their reliance on familiarity as a compensatory mechanism for their recollection deficits. Our results do not support this idea because familiarity declines in the same proportion as recollection during healthy aging and AD.

Third, our results also showed that phonological false recognition increases with age and Alzheimer's disease, as commonly found in the literature (Summer & Huff, 2003; Watson et al., 2001). However our results disagree somewhat with those of Budson et al. (2003), who also found that older adults showed greater levels of relative phonological false recognition than younger adults, but their AD patients exhibited lower levels of phonological false recognition

than older adults. The discrepancy between our results and those of Budson et al. (2003) could be due to the characteristics of the different experimental paradigm used. Budson et al. (2003) used a variation on the DRM phonological false memory paradigm, where each word list is designed to increase activation of a single critical item by presenting phonological neighbors of that item. In contrast, the current paradigm is designed to increase activation of any studied or critical word composed of letters from a subset of the alphabet. In addition, our study lists are significantly longer than those used in the DRM paradigm (34 study words in our case and 12 in the DRM paradigm used by Budson et al., 2003), and this could lead to greater activation of our critical words, which could increase false recognition in older people and, especially, AD patients. The increased incidence of false memories in the latter two groups, according to the activation-monitoring framework (Gallo, 2010; Roediger et al., 2001), would be attributable to their impaired ability to inhibit activation levels on these phonologically associated but non-presented words.

Overall, our results show that our paradigm is sensitive to eliciting the phonological false recognition effect and capturing the impairment in recollection and familiarity and the increase in phonological false memories experienced due to healthy and pathological aging. Therefore, it seems to be an adequate tool for detecting cognitive deterioration in one unique task through the indexes of recollection, familiarity, and relative false recognition. In addition, our paradigm seems to be an innovative way to measure phonological false memories, going beyond the DRM-like approach, which is usually limited to one critical test item per list. Some of the limitations of our study are the small sizes of our samples, especially the sample of AD patients, and the fact that there is no sample of amnestic MCI patients, which would have allowed us to obtain more precise intermediate data on the evolution of the cognitive markers from healthy aging to clinical Alzheimer's disease. Future studies should analyze these ideas.

Acknowledgements

This research was supported by grant PSI2016-77405-R of the Spanish Ministerio de Ciencia, Innovación y Universidades – Agencia Estatal de Investigación (AEI/FEDER, UE).

References

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano* [Frequency Dictionary of Castellan Linguistic units]. Servicio de Publicaciones de la Universidad de Oviedo.

Ally, B. A., Gold, C. A., & Budson, A. E. (2009). An evaluation of recollection and familiarity in Alzheimer's disease and mild cognitive impairment using receiver operating characteristics. *Brain and Cognition, 69*, 504-513. http://dx.doi.org/10.1016/j.bandc.2008.11.003

Basile, B. M., & Hampton, R. R. (2013). Recognition errors suggest fast familiarity and slow recollection in rhesus monkeys. *Learning & Memory, 20,* 431-437. http://dx.doi.org/10.1101/lm.029223.112

Brandt, K. R., Eysenck, M.W., Nielsen, M.K., & von Oertzen, T.J. (2016). Selective lesion to the entorhinal cortex leads to an impairment in familiarity but not recollection. *Brain and Cognition, 104,* 82-92. http://dx.doi.org/10.1016/j.bandc.2016.02.005

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review, 110,* 762-784. http://dx.doi.org/10.1037/0033-295X.110.4.762

Budson, A. E., Sullivan, A. L., Daffner, K. R., & Schacter, D. L. (2003). Semantic versus phonological false recognition in aging and Alzheimer's disease. *Brain and Cognition, 51,* 251-261. http://dx.doi.org/10.1016/S0278-2626(03)00030-7

Devitt, A.L., & Schacter, D.L. (2016). False memories with age: Neural and cognitive underpinnings. *Neuropsychologia, 91*, 346-359. http://dx.doi.org/10.1016/j.neuropsychologia.2016.08.030

Dubois, B., Hampel, H., Feldman, H.H., Scheltens, P., Aisen, P., Andrieu, S., Bakardjian, H., Benali, H., Bertram, L., Blennow, K., Broich, K., Cavedo, E., Crutch, S., Dartigues, J. F., Duyckaerts, C., Epelbaum, S., Frisoni, G. B., Gauthier, S., Genthon, R., ... Cliff, R. J. Jr. (2016). Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. *Alzheimer's & Dementia, 12*, 292-323. http://dx.doi.org/10.1016/j.jalz.2016.02.002

Finley, J. R., Sungkhasettee, V. W., Roediger III, H.L, & Balota, D.A. (2017). Relative contributions of semantic and phonological associates to over-additive false recall in hybrid DRM lists. *Journal of Memory and Language, 93*, 154-168. http://dx.doi.org/10.1016/j.jml.2016.07.006

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189-198. http://dx.doi.org/10.1016/0022-3956(75)90026-6

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition, 38,* 833-848. http://dx.doi.org/10.3758/MC.38.7.833

Hildebrant, H., Haldenwanger, A., & Eling, P. (2009). False recognition helps to distinguish patients with Alzheimer's disease and amnesic aMCI from patients with other kinds of dementia. *Dementia and Geriatric Cognitive Disorders, 28*, 159-167. http://dx.doi.org/10.1159/000235643

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28. http://dx.doi.org/10.1037/0033-2909.114.1.3

King, D. R., de Chastelaine, M., Elward, R. L., Wang, T. H., & Rugg, M. D. (2018). Dissociation between the neural correlates of recollection and familiarity in the striatum and hippocampus: Across-study convergence. *Behavioural Brain Research, 354*, 1-7. http://dx.doi.org/10.1016/j.bbr.2017.07.031

Koen, J.D., & Yonelinas, A.P. (2014). The effects of healthy aging, amnestic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: A Meta-Analytic Review. *Neuropsychology Review, 24,* 332-354. http://dx.doi.org/10.1007/s11065-014-9266-5

Koen, J.D., & Yonelinas, A.P. (2016). Recollection, not familiarity, decreases in healthy ageing: Converging evidence from four estimation methods. *Memory, 24,* 75-88. http://dx.doi.org/10.1080/09658211.2014.985590

McCabe, D. P., Roediger, H. L. III, McDaniel, M. A., & Balota, D. A. (2009). Aging reduces veridical remembering but increases false remembering: Neuropsychological test correlates of remember-know judgements. *Neuropsychologia, 47*, 2164-2173. http://dx.doi.org/10.1016/j.neuropsychologia.2008.11.025

Parkin, A., Ward, J., Squires, E., Furbear, H., Clark, A., & Townshend, J. (2001). Data-driven recognition memory: A new technique and some data on age differences. *Psychonomic Bulletin and Review, 8*, 812-819. http://dx.doi.org/10.3758/BF03196222

Pitarque, A., Sales, A., Meléndez, J.C., Mayordomo, T., & Satorres, E. (2015). An overall decline both in recollection and familiarity in healthy aging. *Psicothema, 27,* 362-367. http://dx.doi.org/ 10.7334/psicothema2015.78

Pitarque, A., Satorres, E., Escudero, J., Algarabel, S., & Meléndez, J.C. (2019). Phonological false recognition produced by bottom-up automatic activation in young and older people. *Memory, 27,* 528-535. http://dx.doi.org/10.1080/09658211.2018.1532520

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1-75. https://doi.org/10.1016/1041-6080(95)90031-4

Rodríguez-Ferreiro, J., Martínez, C., & Cuetos, F. (2019). Differential effects of negative and positive emotional content over veridical and false recognition in aging and Alzheimer's disease. *Journal of Neurolinguistics, 49*, 109-116. http://dx.doi.org/10.1016/j.jneuroling.2018.10.001

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 803-814.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*, 385-407. http://dx.doi.org/10.3758/BF03196177

Scalici, F., Caltagirone, C., & Carlesimo, G. A. (2017). The contribution of different prefrontal cortex regions to recollection and familiarity: A review of fMRI data. *Neuroscience and Biobehavioral Reviews, 83*, 240-251. http://dx.doi.org/10.1016/j.neubiorev.2017.10.017

Schoemaker, D., Gauthier, S., & Pruessner, J.C. (2014). Recollection and familiarity in aging individuals with mild cognitive impairment and Alzheimer's disease: A literature review. *Neuropsychology Review, 24*, 313-331. http://dx.doi.org/10.1007/s11065-014-9265-6

Schoemaker, D., Mascret, C., Collins, D.L., Yu, E., Gauthier, S., & Pruessner, J.C. (2017). Recollection and familiarity in aging individuals: Gaining insight into relationships with medial temporal lobe structural integrity. *Hippocampus, 2017*, 692-701. http://dx.doi.org/10.1002/hipo.22725

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Psychology Software Tools, Inc.

Sommers, M. S., & Huff, L. M. (2003). The effects of age and dementia of the Alzheimer's type on phonological false memories. *Psychology and Aging, 18*, 791-806. http://dx.doi.org/10.1037/0882-7974.18.4.791

Watson, J. M., Balota, D. A., & Roediger, H. L. (2003). Creating false memories with hybrid lists of semantic and phonological associates: Over-additive false memories produced by converging associative networks. *Journal of Memory and Language, 49*, 95-118. http://dx.doi.org/10.1016/S0749-596X(03)00019-6

Watson, J. M., Balota, D. A., & Sergent-Marshall, S. D. (2001). Semantic, phonological, and hybrid veridical and false memories in healthy older adults and in individuals with Dementia of the Alzheimer type. *Neuropsychology, 2*, 254-267. http://dx.doi.org/10.1037/0894-4105.15.2.254

Wolk, D. A., Dunfee, K. L., Dickerson, B. C., Aizenstein, H. J., & DeKosky, S. T. (2011). A medial temporal lobe division of labor: insights from memory in aging and early Alzheimer disease. *Hippocampus, 21*, 461-466. http://dx.doi.org/10.1002/hipo.20779

Wolk, D.A., Mancuso, L., Kliot, D., Arnold, S.E., & Dickerson, D.C. (2013). Familiarity-based memory as an early cognitive marker of preclinical and prodromal AD. *Neuropsychologia, 51*, 1094-1102. http://dx.doi.org/10.1016/j.neuropsychologia.2013.02.014