Published in final edited form as:
Brain Lang. 2021 September ; 220: 104982. doi:10.1016/j.bandl.2021.104982.

Age-related Differences in the Retrieval of Phonologically Similar Words during Sentence Processing: Evidence from ERPs

Hossein Karimi[1], **Michele Diaz**[2]

[1]Mississippi State University

[2]The Pennsylvania State University

Abstract

We investigated how phonologically similar words are encoded and retrieved from memory during sentence processing across younger and older adults. Critical sentences included two phonologically similar or dissimilar noun phrases (henceforth NPs) followed by a pronoun. We examined brain activity time-locked to the onsets of the second NP, and the pronoun to investigate the encoding and retrieval of the NPs, respectively. Encoding the second NP resulted in smaller N400 amplitudes when the preceding NP was phonologically similar, for both younger and older adults, suggesting age-invariant encoding facilitation with increasing phonological similarity. However, when processing the pronoun, younger adults exhibited greater negativity following phonologically similar NPs, suggesting retrieval difficulty, whereas older adults showed greater negativity for pronouns following *dis*similar NPs, suggesting an apparent retrieval facilitation. A post-hoc behavioral experiment suggested that older adults perform shallow processing during retrieval. The results suggest age-related decline in retrieval, but not encoding, of phonological information.

Keywords

Aging; Language processing; Encoding; Retrieval; Phonological similarity

1. Introduction

Language processing necessarily involves encoding, storage and retrieval of information from memory so that a coherent discourse representation can be formed. Numerous previous studies have empirically shown that general memory functions such as encoding, storage, and retrieval support language processing (Hofmeister, 2011; Jäger et al., 2017; Lewis et al., 2006; Lewis & Vasishth, 2005; McElree, 2000; McElree et al., 2003; Patil et al., 2016; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006). A prime example of these memory processes is referential processing, whereby one or multiple noun phrases (NPs) are

Corresponding author: Hossein Karimi, Department of Psychology, Mississippi State University, 215 Magruder Hall, 255 Lee Boulevard, Mississippi State, MS 39762, h.karimi@msstate.edu, (662) 325-3202.

encoded in memory and then retrieved at a subsequent point. For example, in a discourse segment such as *John waved to Sara from the other side of the room when she was having a coffee,* memory representations for the involved human entities (i.e., *John* and *Sara*) need to be formed and encoded in memory (along with representations of the other entities and the relations among them). Then, when a subsequent pronominal reference is made to *Sara* by the pronoun *she*, the memory representation of the NP associated with the pronoun (henceforth, the *referent*) needs to be reactivated and retrieved from memory so that the referential dependency can be resolved (Dell, McKoon, & Ratcliff, 1983; Gernsbacher, 1989; Gerrig & McKoon, 1998; Lucas, Tanenhaus, & Carlson, 1990; MacDonald & MacWhinney, 1990; Sanford & Garrod, 1989; Sanford & Garrod, 2005).

A number of factors have been shown to influence retrieval, including semantic similarity between referential candidates, and discourse-related factors such as the probability and the number of potential referents. For example, both syntactic and referential dependencies have been shown to be more difficult to process when the relevant NPs are semantically more similar (*Syntactic*: Gordon et al., 2001, 2002, 2006; also see Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006; *Referential*: Patil et al., 2016; Jäger et al., 2015; also see Karimi et al., 2018). In addition, pronominal ambiguity and the probability of referential interpretations have also been shown to influence referential processing, with ambiguous and more equi-probable pronouns causing more processing difficulty than unambiguous and non-equi-probable pronouns, respectively (Nieuwland & Van Berkum, 2006; van Berkum et al., 1999; Van Berkum et al., 2007); an effect that has been attributed to having to entertain one vs. multiple referential interpretations (Nieuwland & Van Berkum, 2006). Collectively, these studies show a clear effect of semantic and discourse factors on referential processing.

A family of psycholinguistic theories commonly known as cue-based retrieval theories propose a mechanism underlying long-distance linguistic dependency resolutions such as syntactic and referential dependencies (Jäger et al., 2017; Lewis et al., 2006; Lewis & Vasishth, 2005). Based on these accounts, a pronoun sets morpho-syntactic retrieval cues such as gender, number, etc., and these cues trigger the retrieval of the felicitous referent via a content-addressable memory probe (see McElree, 2000; McElree et al., 2003). Crucially, under cue-based theories, the amount of interference experienced during the retrieval of an item in memory affects the retrieval ease of that item (Lewis et al., 2006; Lewis & Vasishth, 2005), explaining the semantic similarity and the discourse effects mentioned above. Although prior studies have identified semantic and discourse factors that influence the encoding and retrieval of words during sentence processing, the roles of phonology and age-related cognitive decline remain under-investigated. Understanding how phonology affects retrieval is of particular importance to theories of aging and language as current accounts have suggested that age-related language decline may be phonologically based.

1.1. The role of phonology

An important assumption of cue-based retrieval theories is that access to stored representations is made possible through semantic and/or morpho-syntactic features (such as gender, animacy, humanness, singularity, etc.), but not through phonological cues. Although it is also possible that phonological information could be used as a cue, most current cue-based retrieval theories do not make explicit claims about the role of phonology during retrieval (see Jäger et al., 2017; Lewis et al., 2006; Lewis & Vasishth, 2005, among many others). Intrinsic to these models is the idea of "cues" as guides to successful retrieval. Although morpho-syntactic features such as gender (e.g., [+feminine]), and number (e.g., [+singular]) can straightforwardly constitute retrieval cues for specific words, it is not clear how phonological features would constitute retrieval cues. For example, phonological features such as, say, [+voiced], and [+bilabial] are associated with individual phonemes rather than words. One possibility is to propose phoneme-based retrieval cues such as [+has a "b" sound]. However, current cue-based are underspecified with regards to the potential role of phonological retrieval cues. Importantly, these theories assume that attentional resources are very limited, and once incoming linguistic information is encoded, it is rapidly transmitted to the long-term memory store, access to which is made possible through semantic and/or morpho-syntactic codes. Thus, phonological information may affect how incoming words are *encoded*, with shared phonological features leading to encoding interference, but not how they are *retrieved* (Kush et al., 2015).

An alternative account for the potential effect of phonology during sentence processing comes from theories that assume a phonologically-mediated working memory store, which actively maintains the phonological codes of words during sentence processing (Caramazza et al., 1981; Gibson, 1998; Shankweiler & Crain, 1986). These theories are in line with findings from the memory literature showing that word lists with phonological overlap are more difficult to recall compared with word lists without phonological overlap (e.g., Baddeley, 1966, 2018; Craik, 1968). Thus, under these accounts, phonological similarity should produce interference during both encoding and retrieval of word representations during sentence processing.

Interestingly, both accounts have received support from prior experimental work. For instance, using sentences such as (1), Acheson and MacDonald (2011) showed that sentence processing is rendered more difficult when phonological overlap between words is high (e.g., *baker* and *banker*, and *sought* and *bought* in 1a), than when it is low (e.g., *runner* and *banker*, and *feared* and *bought* in 1b). The critical issue regarding sentences such as (1) is that successful processing of the embedded verb (i.e., *sought* in 1a, or *feared* in 1b) depends on the retrieval of its syntactic object (i.e., *baker* in 1a, and *runner* in 1b) that has been moved from its canonical position (i.e., immediately following the embedded verb) to an earlier position in the sentence to create emphasis (e.g., Fodor, 1978; Hofmeister, 2011; Kluender & Kutas, 1993; McElree, 2000; Nicol & Swinney, 1989; Osterhout & Swinney, 1993). Thus, greater difficulty processing (1a) relative to (1b) suggests that the phonological codes of words are actively maintained in memory and complicate the encoding and/or retrieval of memory items, and therefore sentence processing in general. Similarly, Nakayama, Lee, & Lewis (2005) showed that morpho-phonological similarity between linguistic memory items produce interference during retrieval in Korean, supporting the operation of "similarity-based interference in working memory" (also see Lewis & Nakayama, 2002; and Vasishth, 2002).

a. *The baker that the banker sought bought the house.*

b. *The runner that the banker feared bought the house.*

More recently, a research study by (Kush et al., 2015) investigated whether phonological overlap causes interference during encoding, retrieval, or both processes. Participants read sentences such as (2) while maintaining three words in memory.

(2) It was the boat that the guy who drank some hot coffee sailed on two sunny days.—Similar to (1), successful processing of the verb *sailed* in (2) depends on the retrieval of its syntactic object (i.e., *boat*; e.g., Fodor, 1978; Hofmeister, 2011; Kluender & Kutas, 1993; McElree, 2000; Nicol & Swinney, 1989; Osterhout & Swinney, 1993)[1]. Kush et al., (2015) manipulated the phonological similarity of a three-word memory load such that they were either phonologically similar to the to-be-retrieved *boat* representation (*coat–vote–note*) or not (*table–sink–truck*). The researchers observed longer reading times for *boat* when the memory load items sounded similar to it, indicating an encoding interference effect. However, reading times did not reliably vary as a function of phonological similarity during retrieval, namely, when *sailed* was being read/processed. Based on these results, the authors argued that although phonological similarity might cause interference during encoding, it does not affect the retrieval of associated memory items, lending support to the claim by cue-based theories that phonological features do not produce interference during retrieval[2].

While Kush et al., (2015)'s experiments were well-suited for testing the effect of memory processes on sentence comprehension, at least two main issues could be identified with their design. First, natural language processing does not normally involve maintaining other unrelated words in memory. More specifically, although manipulating the memory load items might have influenced the encoding and retrieval of the critical words, they were not intended to be *integrated* into the sentence representation (unlike what happens during natural sentence processing). Second, self-paced reading is limited to measuring reaction times and does not allow a more nuanced examination of encoding and retrieval operations. In fact, the encoding interference effect reported by Kush et al., (2015) runs counter to numerous studies reporting facilitated lexical access due to phonological overlap (Carreiras et al., 2005; Humphreys et al., 1982; Lukatela & Turvey, 1994). Thus, although prior studies provide some insight into how sentence processing, and encoding and retrieval operations might be affected by phonological similarity, the evidence is currently not strong enough to permit firm conclusions.

1.2. The role of cognitive decline

Another important factor that has been shown to influence phonological processing is cognitive aging. Specifically, it has been shown that although older adults have largely intact semantic processes, they experience considerable difficulty with phonological aspects of word retrieval. For instance, older adults have been reported to be slower and less

[1]The "normal" (not clefted) construction of (2) would be: *The guy who drank some hot coffee sailed the boat on two sunny days*. It-cleft constructions put a semantic emphasis on the target object (in this case *boat*) by moving it to an earlier position in the sentence.
[2]There is evidence in the literature that the syntactic class of words may modulate the effect of phonological similarity, with same-class, phonologically similar primes having no effect on recovery from Tip of the Tongue state, but different-class primes facilitating recovery (Abrams & Rodriguez, 2005). Although these results reflect phonological priming during language *production* (not comprehension), because our target name pairs were always restricted to the same syntactic class (i.e., nouns; see below), our results may only generalize to this syntactic class.

accurate during picture naming (Diaz et al., 2014, 2019, 2021; Rizio et al., 2017; Taylor & Burke, 2002; Zhang et al., 2019), experience more tip-of-the-tongue states where the meaning of a target word is known but its phonological code cannot be fully retrieved (R. Brown & McNeill, 1966; Burke et al., 1991; Cross & Burke, 2004), and produce more pauses during speech (Kemper et al., 1992). One potential explanation for such cognitive decline comes from the Transmission Deficit Hypothesis (e.g., Burke et al., 1991, 2000). Based on this account, while aging negatively affects all connections, phonological representations are more vulnerable to decline. This is because the semantic network is highly interconnected whereas the phonological network is more sparsely connected. As a result, when connections undergo age-related decline, semantic processes are affected less than phonological processes[3]. Consistent with this account, experimental work has demonstrated greater language production difficulty for older compared to younger adults, for phonological, compared with sematic information (e.g., Diaz et al., 2014; MacKay & James, 2004; Taylor & Burke, 2002). While the majority of these deficits have been observed during language production, collectively these findings suggest age-related differences in phonological retrieval. This is consistent with the wider memory literature that has demonstrated age-related declines in encoding and retrieval of information from memory (e.g., Bowles & Poon, 1985; Craik & Rose, 2012). Thus, based on these prior literatures, we would expect older adults to exhibit more difficulty retrieving phonologically similar words during sentence processing.

1.3. The present study

The aim of the present study is to investigate the role of phonology during the encoding and retrieval of referential candidates, and to investigate the potential interaction between phonological processing and cognitive aging. We created experimental stimuli such as (3) in which two phonologically similar or dissimilar NPs (in the form of proper names) had to be encoded sequentially, followed by a pronoun that either unambiguously referred to one of the NPs (i.e., the referential candidates), or was formally ambiguous (i.e., could refer to either NP). Thus, the design crossed Phonological Similarity of the referential candidates (Similar vs. Dissimilar) with Gender Congruence/Pronominal Ambiguity (Same Gender/Ambiguous vs. Different Genders/Unambiguous)[4].

(3). Sample Experimental stimuli.

3a)	Phonologically Similar Same Gender:	Jason laughed at Jacob when he was almost drunk and high.
3b)	Phonologically Similar Different Genders:	Jade laughed at Jacob when he was almost drunk and high.
3c)	Phonologically Dissimilar Same Gender:	Matt laughed at Jacob when he was almost drunk and high.

[3]Note that although aging is associated with decline in some fundamental cognitive processes such as processing speed and episodic memory, because some language-related processes such as semantic processing remain virtually intact, the term "age-related cognitive *differences*" or "age-related cognitive *changes*" (rather than "*decline*") might better capture the effect of aging in the realm of language. However, because we focus on phonological processing, and because phonological processing mostly declines with age, we will continue to use the term "cognitive decline" in this paper.

[4]Because the gender congruence of the names determines the ambiguity of the pronoun, Gender Congruence (Same vs. Different-genders) and Pronominal Ambiguity (Ambiguous vs. Unambiguous) are essentially the same manipulation. However, in order to maximize clarity, we will refer to this manipulation as Gender Congruence when analyzing how NP2 was encoded after processing NP1, and as Pronominal Ambiguity when analyzing how the pronoun was processed (see below).

3d)	Phonologically Dissimilar Different Genders:	Hannah laughed at Jacob when he was almost drunk and high.

We then examined electrophysiological brain activity time-locked to two points in the critical sentences: The onset of the second name (i.e., NP2), and, the onset of the pronoun. Brain activity at the onset of NP2 would reveal if phonological similarity affects the encoding of words during sentence processing, and brain activity during processing the pronoun would show whether the retrieval of referential candidates is sensitive to phonological similarity. In addition to brain activity, we also conducted a separate behavioral experiment where we directly assessed how pronouns were interpreted as a function of Age and Phonological Similarity.

1.4. Predictions

Since the N400 component is sensitive to integration difficulty (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980),we hypothesized that ease of encoding of NP2 should cause variations in this component. Specifically, when encoding the second NP, if phonological similarity causes encoding interference, NP2s should yield larger N400 amplitudes when the preceding NP1s are phonologically similar than when they are dissimilar (Kush et al., 2015, also see Caramazza et al., 1981; Gibson, 1998; Shankweiler & Crain, 1986). Alternatively, if phonological similarity leads to phonological priming and facilitation (Carreiras et al., 2005; Desroches et al., 2009; Humphreys et al., 1982; Lukatela & Turvey, 1994), we should observe smaller N400 amplitudes for NP2s following phonologically similar compared to phonologically dissimilar NP1s.

When processing the pronoun (i.e., when the memory representation associated with the referent is being retrieved), if phonological similarity does not influence retrieval, we would expect a null effect of phonological similarity on pronoun processing (Kush et al., 2015; Lewis et al., 2006; Lewis & Vasishth, 2005). However, if working memory is phonologically mediated, phonological similarity should complicate the retrieval of the associated representations. But how would this difficulty manifest itself? One possibility is an Nref effect (van Berkum et al., 1999). The Nref is a sustained and frontal negativity elicited for ambiguous vs. unambiguous pronouns (van Berkum et al., 1999), and also reflects retrieval difficulty (Karimi et al., 2018). However, another possibility is that phonological similarity might lead to a brain response associated with memory recognition difficulty. Specifically, past research into recognition memory has reported a late posterior negativity (henceforth LPN) as a function of correctly retrieving contextual information, and continued evaluation of retrieval outputs (see Johansson & Mecklinger, 2003, and Mecklinger et al., 2016 for a review). Moreover, past research into referential processing has also demonstrated the engagement of recognition memory processes during referential processing (Nieuwland & Martin, 2017). Thus, to the extent that phonological similarity engages recognition memory processes, we might expect an LPN effect for successful phonological retrieval. Note that, regardless of phonological similarity of the two NPs, we should also observe an Nref effect for ambiguous relative to unambiguous pronouns (e.g., Nieuwland, 2014; Nieuwland & Van Berkum, 2006; van Berkum et al., 1999; Van Berkum et al., 2007).

Finally, with regards to the potential effect of age-related cognitive decline, we expect older adults to show similar patterns of brain activity as younger adults during *encoding,* as older adults often do not show large age-related declines in single word reading. However, as our discussion of the prior literature has shown (as discussed above, Burke et al., 1991, 2000; Diaz et al., 2014), older adults should exhibit more *retrieval* difficulty than younger adults when processing the pronoun. Such greater retrieval difficulty might translate into larger Nref or LPN effects than younger adults. However, the relationship between task difficulty in general and retrieval difficulty in particular, and brain response might not be linear. In fact, previous research has shown that at higher levels of task difficulty, older adults tend to have poorer performance and engage fewer brain regions (Reuter-Lorenz & Cappell, 2008; Zhang et al., 2019). These results are in line with the Good Enough view of language processing. Based on this approach, language processing might rely on “heuristics”, or “shallow” processing to reduce the amount of attentional and/or memory resources devoted to the task at hand. Critically, in cases where processing can proceed without resolving a linguistic dependency, that dependency might be left unspecified (Ferreira, 2003; Ferreira et al., 2002; Karimi & Ferreira, 2016a; Sanford & Sturt, 2002; Stewart, Holler, & Kidd, 2007; Swets et al., 2008). Thus, an alternative possibility is that if processing phonological information is too difficult for older adults, we might observe a type of good-enough processing where the processing difficulty is avoided by shallow processing, perhaps due to lack of sufficient attentional and/or memory resources. Evidence for such shallow processing may come from an experiment directly assessing whether and how older adults resolve the pronoun. If phonological similarity causes older adults to resort to shallow processing during pronoun resolution (i.e., at retrieval), they should exhibit lower accuracy for comprehension questions directly probing the referent of the pronoun.

2. EEG Experiment: Methods

2.1. Participants

Forty-seven younger adults, and 41 older adults participated in the experiment. However, data from 7 younger participants and 1 older participant were removed from the analyses due to excessive noise, leaving 40 participants in each group (Younger adults: Age range: 18–29, mean age = 19.37, 31 females, 9 males; Older adults: Age range: 59–76, mean age = 67.22, 26 females, 14 males). Participants were paid $10/hour or received course credit for their participation. All participants were right-handed, monolingual, native speakers of American English, who reported no neurological or language-related disorders, and had no exposure to any other languages before the age of 5. Additionally, all participants took the Mini-Mental State Exam test (Folstein et al., 1975), limiting the possibility of mild cognitive impairment/dementia (mean = 28.92, SE = 1.05, range = 27–30). All participants provided informed consent and all experimental procedures were reviewed and approved by the Institutional Review Board at the Pennsylvania State University.

2.2. Stimuli

We used a total of 100 experimental stimuli such as (3), yielding 25 items per condition.

(3). Sample Experimental stimuli.

3a)	Phonologically Similar-Same Gender:	Jason laughed at Jacob when he was almost drunk and high.
3b)	Phonologically Similar-Different Genders:	Jade laughed at Jacob when he was almost drunk and high.
3c)	Phonologically Dissimilar-Same Gender:	Matt laughed at Jacob when he was almost drunk and high.
3d)	Phonologically Dissimilar-Different Genders:	Hannah laughed at Jacob when he was almost drunk and high.

Because it was difficult to find names that were frequent, familiar, and shared phonological features in the same location, we selected the critical names such that they had *onset* overlap in half of experimental items (e.g., *Jacob and Jason/Jade* vs. *Jacob and Matt/Hannah*), and *rhyme* overlap in the other half (e.g., *Irvin/Erin and Kelvin* vs. *Curtis/Mackenzie and Kelvin*), as illustrated in (3). However, because we did not find any reliable differences as a function of overlap locus, we collapsed over this variable for all the analyses and graphs (see below)[5]. It is important to mention that there was a good deal of variability in terms of the amount of phonological overlap between the two names across our items. However, the overall phonological similarity between the critical Phonologically Similar and Phonologically Dissimilar conditions was significantly different, and this difference did not stem from a subset of items (see the "Phonological Similarity Check" section below). Moreover, the number of words intervening between NP2 and the critical pronoun ranged from one to four words. The full list of experimental stimuli is provided as supplementary materials available at: https://osf.io/75qyz/. It is important to note that we intentionally chose proper names (rather than common nouns such as *pilot* and *pirate*) to better isolate the processing of phonological information, and because older adults have been shown to experience more difficulty retrieving proper names relative to common nouns (Cohen & Faulkner, 1986; Lovelace & Twohig, 1990; Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014). Thus, using names would maximize the power of our study. Also note that, although our experimental stimuli were single sentences, they were natural English sentences, easily processed by native speakers and consistent with prior studies (e.g., Nieuwland & Van Berkum, 2006; Nieuwland et al., 2007; van Berkum et al., 1999).

The sentences were presented to the participants along with 60 fillers. Twenty-four of the critical sentences and 20 of the fillers were tagged with a comprehension question in the form of a True/False statement. The questions always assessed the general meaning of the sentence, to ensure that participants were paying attention (e.g., *Jason laughed at Jacob*). Note that we did not directly probe the referent of the pronouns because previous research has shown that such questions may alter the strategies people adopt for sentence processing (Stewart, Holler, & Kidd, 2007; Swets et al., 2008), and could have, therefore, altered brain activity during pronoun processing. However, we did run a follow-up behavioral experiment to directly assess pronoun resolution across all items (see below). Importantly, the comprehension questions during the EEG experiments were not the same for all versions of an experimental sentence; some of the questions varied based on

[5]The effects (see below) were slightly stronger for onset-overlapping names compared to rhyme-overlapping names, which is consistent with previous research (Fricke et al., 2016). However, the direction of the effects were the same across both overlap types.

Gender Congruence/Pronominal Ambiguity. This was because unlike ambiguous pronouns, unambiguous pronouns permitted asking more specific questions (e.g., *Hannah was drunk*).

We counterbalanced the linear positions of the NPs across participants, in order to limit any potential effects of frequency, length, or familiarity of the names (i.e., that a pronoun could be taken as referring to a particular NP only because that name is more frequent/familiar/ etc.). In the context of (3), this means that for half of the items *Jacob* was the second NP and for half of the experimental stimuli, it was the first NP. As a result of this, we ended up with 8 experimental lists (4 lists capturing the 4 experimental conditions and 4 lists capturing the same conditions, but in which the linear order of the names was reversed). Each critical item appeared only once in each list, thus avoiding potential repetition effects. The items in each list were randomized for each participant, so that each participant read the stimuli in a different order. The experiment was programmed in E-prime 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA), which displayed the stimuli and recorded codes time-locked to the onset of critical NPs (i.e., NP2s), and pronouns in each sentence.

Importantly, before running the EEG experiment, we assessed two important characteristics of our stimuli. First, since past research has shown that the probability of referential interpretations affect ease of pronoun processing (e.g., Nieuwland & Van Berkum, 2006), we first conducted a norming study with a separate group of participants to measure the probability of resolving the pronoun as referring to NP1 vs. NP2 across the two ambiguous conditions. Second, to ensure that our Phonological Similarity manipulation actually worked, we calculated the degree of phonological similarity between the two NPs across the Phonologically Similar and Dissimilar conditions.

2.3. Norming Experiment—Following previous research (Koornneef & Vanberkum, 2006; Nieuwland & Van Berkum, 2006; Van Berkum et al., 2007), we presented the participants with sentence fragments including all the words up to the critical pronoun, resulting in discourse fragments such as (4). Twenty-four younger participants (mean age = 19.95, age range = 18–22, 17 females, 7 males), and 26 older adults (mean age = 67.15, age range = 60–74, 16 females, 10 males) were asked to provide a continuation for each sentence fragment and then explicitly indicate which referential candidate they talked about in their continuation; NP1 (e.g. *Jason*/*Matt*) or NP2 (e.g., *Jacob*). None of the participants took part in the main EEG experiment.

4a)	Phonologically Similar-Same Gender/Ambiguous:	Jason laughed at Jacob when he ...
4b)	Phonologically Dissimilar-Same Gender/Ambiguous:	Matt laughed at Jacob when he ...

The overall tendency to interpret the pronoun as referring to NP1 vs. NP2 was virtually the same across conditions for younger adults (Phonologically Similar: NP1 interpretation = 51.7%, NP2 interpretation = 48.3%; Phonologically Dissimilar: NP1 interpretation = 49.6%, NP2 interpretation = 50.4%), and for older adults (Phonologically Similar: NP1 interpretation = 58.8%, NP2 interpretation = 41.2%; Phonologically Dissimilar: NP1 interpretation = 58%, NP2 interpretation = 42%). We ran a mixed-effects logistic regression model using the lme4 package in the statistical software R (version 3.6.1), with Age and Phonological Similarity as well as their interaction as predictors. The random effects structure of the model included intercepts for both subjects and items, by-subjects slopes for the effect of Phonological Similarity, and by-item slopes for the effects of Age, Phonological Similarity, and their interaction (Barr et al., 2013). The results of this regression model showed no effect of Age ($\beta = .54$, $SE = .49$, $z = 1.10$, $p = .27$), no effect of Phonological Similarity ($\beta = -.07$, $SE = .08$, $z = -.87$, $p = .38$), and no interaction ($\beta = -.08$, $SE = .14$, $z = -.61$, $p = .54$). These results suggest that the choice between NP1 and NP2 was not affected by phonological similarity for either group.

2.4. Phonological Similarity Check

To test whether the two NPs in the phonologically similar condition actually sounded more similar compared to the two NPs in the dissimilar condition, we calculated the Levenshtein distance between the *pronunciation* of the two NPs using the Carnegie-Mellon Pronouncing Dictionary, version 0.7b (http://www.speech.cs.cmu.edu/cgi-bin/cmudict) and the *adist* function of R. Thirty-six names (out of 500 total names[6]) did not have corresponding pronunciation entries in the Carnegie-Mellon Pronouncing Dictionary. The mean Levenshtein distance between the phonology of NPs in the Phonologically Similar and Dissimilar conditions were 5.68 and 10.3, respectively. Note that *smaller* distances denote *more* similarity. We then ran a mixed-effects regression model on the Levenshtein distances with random intercepts for items and by-item random slopes for the effect of Phonological Similarity to ensure that any significant difference did not arise from a subset of items. This model revealed a statistically significant difference in Phonological Similarity, ($\beta = -4.58$, $SE = .29$, $z = -15.61$, $p = < .001$). We provide the Levenshtein distance values between the two names for all of our items, along with the R analysis script in the supplementary materials (https://osf.io/75qyz/).

2.5. Procedure

Before starting the EEG recording session, each participant took an online version of the Reading Span test (Daneman & Carpenter, 1980; Loboda, 2012), the Author Recognition Test (Stanovich & West, 1989), and the vocabulary portion of the Nelson-Denny test (Brown, 1960). These cognitive tests were measured to examine correlations with the brain responses (see below). During the EEG recording session, participants were seated in front of an LCD computer screen placed approximately 90 cm in front of them. Sentences were presented on the screen one word at a time, and participants were asked to read them for comprehension. Each word stayed on the screen for 200 ms, followed by 300 ms of blank screen. The sentences were always preceded by a fixation cross to direct the participants to look at the middle of the screen (where the words were displayed) and to reduce eye-movement related artifacts. After 27.5% sentences, a comprehension question appeared on the screen, and the participants were instructed to indicate whether the sentence was true or false through a button press. They had unlimited time to answer the question. After the comprehension question was answered, or when the presentation of a sentence was over (for sentences without a question), the participants pressed another button to start the next trial. The participants were instructed to read the sentences carefully

[6]Note that each experimental item consists of 5 unique names. See (3) above.

for comprehension and to do their best to answer the questions as accurately as possible. To minimize physiological artifacts on the EEG signal, the participants were instructed to sit as still as possible and refrain from blinking and/or moving their eyes during the presentation of each trial. They were told they could blink between trials and also when answering comprehension questions. The experimental session was divided into four blocks. A practice session preceded the main experiment and consisted of 6 trials (3 of which included comprehension questions), which allowed participants to become accustomed to the experiment. The experiment took approximately 40 minutes for younger adults and 55 minutes for older adults to complete.

2.6. EEG recording

We recorded the electroencephalogram (EEG) from 32 Ag/AgCl electrodes mounted in an elastic cap (Brain Vision acticap). Two of the electrodes were attached below and to the side of each eye to record blinks and horizontal eye movements for later artifact rejection, and another two electrodes were attached to the left and right mastoids for subsequent offline re-referencing. All electrode impedances were kept below 20 kΩ. The EEG signal was digitally recorded at a sampling rate of 500 Hz and was amplified by means of a 32 channel ActiCHamp system (Brain Products, Inc). Preprocessing was done using the EEGlab and ERPlab tools in MATLAB (Mathworks Inc., Natick, MA). Specifically, the recoded EEG data from each participant was bandpass filtered (.1–30 Hz)[7], epoched (see below), and submitted to an independent component analysis (ICA) to isolate and remove EEG components associated with eye blinks. Single-trial waveforms were then screened for other artifacts such as amplifier drift, muscle artifacts, and eye movements. Any epochs containing these artifacts were rejected prior to analysis (approximately 2% of all trials). Then, the data were re-referenced to the average of the left and right mastoids off-line and ERPs were calculated by averaging individual EEG epochs, time-locked to the presentation of the NP2 and the critical pronoun in all 4 conditions. The length of the ERP epochs differed based on whether the corresponding analysis examined encoding of NP2 or retrieval of referential candidates (i.e., at the onset of the critical pronoun). Epochs for encoding NP2 extended from 300 ms before to 800 ms after the presentation of NP2. Epochs for retrieval of the referential candidates extended from 300 ms before until 1500 ms after pronoun onset. For both epochs, the 300 ms preceding the onset of the critical words constituted the baseline. All epochs were baseline-corrected.

2.7. Statistical analyses

In addition to Age as the between-subjects and Phonological Similarity and Pronominal Ambiguity as the within-subjects factors, we entered Brain Region as an additional 3-level within-subjects factor to capture the topographic distribution of the effects. This topographic factor captured the lateral distribution of the effects along the anterior to posterior parts of the scalp with three levels: Frontal (FP1/2, F3/4, F7/8, and Fz), Central (FC1/2, FC5/6, C3/4, CP1/2, CP5/6, and Cz) and Posterior (T5/6, P3/4, O1/2, Pz, and Oz). Thus, when crossed with the main factors of interest, this analysis resulted in a

[7]Note that a high-pass threshold of .1 and smaller thresholds (e.g., .01) have been shown to produce virtually identical results (Tanner, Morgan- Short, & Luck, 2015).

2×2×2×3 mixed effects design: Age (Younger vs. Older) × Phonological Similarity (Similar vs. Dissimilar) × Gender Congruence/Pronominal Ambiguity (Same Gender/Ambiguous vs. Different Genders/Unambiguous) × Anteriority (Frontal vs. Central vs. Posterior).

We always report the main effects of Age, Phonological Similarity, and Gender congruence/Pronominal Ambiguity, as well as their interactions. However, for brevity, any interactions between the topographic regions and the main factors of interest are reported only when they were significant. Any unreported effects were statistically non-significant. Full results along with the raw data are provided as supplementary materials, available at: https://osf.io/75qyz/.

2.8. Behavioral Results (to comprehension questions during the EEG study)

We analyzed both accuracy and reaction times (RTs) to the comprehension questions. As mentioned above, the comprehension questions varied for ambiguous and unambiguous items, which means that the difficulty and the length of the comprehension questions differed based on Gender Congruence/Pronominal Ambiguity, making it impossible to assess the potential effect of this variable on the behavioral results. However, we analyzed the effect of all three predictors (Age, Phonological Similarity and Gender Congruence/Pronominal Ambiguity) and their interactions for maximum transparency. For brevity, we only report the statistically significant results here in this section. However, the full results along with R analysis scripts are provided in the supplementary materials.

2.8.1. Accuracy—The results of a generalized mixed-effects regression model with random intercepts for subjects, by-item random slopes for the interaction term between Age and Gender Congruence/Pronominal Ambiguity, revealed a main effect of Age, with older adults being less accurate than younger adults (Mean accuracy for younger adults = 94.6%, Mean accuracy for older adults = 91.4%, $\beta = -.58$, $SE = .27$, $z = -2.13$, $p = .03$), a main effect of Gender Congruence/Pronominal Ambiguity in the unexpected direction, with Same-Gender/Ambiguous sentences resulting in *more* accuracy (Mean accuracy for Same-Gender/Ambiguous = 94.3%, Mean accuracy for Different-Gender/Unambiguous = 91.6%, $\beta = .55$, $SE = .20$, $z = 2.65$, $p = .008$). Importantly, however, the effect of Gender Congruence/Pronominal Ambiguity depended on Age ($\beta = -1.32$, $SE = .41$, $z = -3.19$, $p = .001$), such that younger adults were more accurate than older adults in the Same-Gender/Ambiguous condition ($\beta = -1.42$, $SE = .70$, $z = -2.02$, $p = .04$), but not in the Different-Gender/Unambiguous condition ($\beta = -.07$, $SE = .37$, $z = -.19$, $p = .84$). Critically, the main effect of Phonological Similarity and the Age × Phonological Similarity interaction were not statistically reliable.

2.8.2. Reaction times—We first cleaned this data by removing any reaction times faster than 1000 ms and slower than 10000 ms, and then removing reaction times that fell 2.5 standard deviations above or below the mean for each subject, which resulted in loss of 73 out of 1920 data points (3.8%). We ran a mixed effect regression model with the full random effects structure (Barr et al., 2013) on log-transformed reaction times (RTs) to the comprehension questions. The results showed a significant main effect of Age, with older adults being slower than younger adults (Mean RT for younger adults = 2817 ms, Mean RT for older adults = 3137 ms, $\beta = .11$, $SE = .05$, $z = 2.34$, $p = .02$). However, the effects of Phonological Similarity and its interaction with Age were not statistically significant.

2.9. Discussion of behavioral results (during EEG)

Although the overall accuracy for both groups was high (> 90%), the behavioral results showed a main effect of age, such that older adults were both less accurate and slower than younger adults when answering the comprehension questions. However, neither the effect of phonologically similarity nor its interaction with age were statistically reliable. It is important to note that because the questions assessed general understanding of the sentences and not pronoun resolution, the behavioral results do not directly speak to pronoun resolution, which relies on successful retrieval of the memory representations associated with the referential candidates (i.e., the NPs).

2.10. EEG Results

We analyzed the data during the encoding and retrieval of referential candidates in 2 separate analyses. In the 'encoding' analysis, we examined the data in the 300–500 ms window following the second NP, and in the 'retrieval' analysis, we analyzed the data in the 700–1100 ms following the critical pronoun. The results of our two analyses are reported below.

2.10.1. Encoding results (300 – 500 ms following NP2 onset)—Figure 1 shows the grand average brain waveforms of younger (Panel A) and older adults (Panel B) during encoding of NP2 for Phonologically Similar and Dissimilar conditions[8]. Following previous research, we limited the analyses to the time window between 300–500 ms after the onset of NP2 (highlighted in Figure 1), where the N400 amplitude is expected to be maximal (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). Table 1 reports the results of our statistical analysis. As can be seen in Table 1, Phonological Similarity interacted with Brain Region such that NP2s resulted in smaller N400 amplitudes when they followed similar-sounding than dissimilar-sounding NP1s at frontal brain regions, but not at central or parietal brain regions (compare the average of two blue lines with the average of the two red lines, moving from frontal to parietal sites). Moreover, we observed a significant main effect of Age, as well as significant Age × Brain Region interaction such that N400 amplitudes were reliably greater for younger relative to older adults at frontal and central, but not parietal brain regions.

2.10.2. Discussion of encoding results—During encoding NP2, both younger and older adults exhibited facilitation as a function of phonological similarity. Specifically, NP2s following similar-sounding NP1s resulted in smaller N400 amplitudes relative to NPs following dissimilar-sounding NP1s at the frontal brain regions. Thus, while there was a significant main effect of Age, the interaction between Age and Phonological Similarity was not significant, suggesting that although there are age-related differences in the overall neural response, older adults were similarly sensitive to phonology during encoding. The facilitatory frontal effect at encoding is consistent with the phonological priming literature

[8]Because Gender Congruence did not have an effect, we are only showing the Phonological Similarity effect for simplicity. However, the plot for all 4 conditions is provided in the Supplementary Materials.

reporting easier processing of verbal stimuli/words as a function of phonological overlap between primes and targets (Desroches et al., 2009; Humphreys et al., 1982), and runs counter to the encoding interference effect reported by Kush et al. (2015). Note that the frontal distribution of this effect might raise concerns about its functional interpretation. However, numerous previous studies have reported frontally distributed N400 effects as a function of conceptual priming (e.g., Voss et al., 2010; Voss & Paller, 2006), as well as phonological priming (Jescheniak et al., 2002). There is some debate on whether the posterior- and frontal-maximal N400 are functionally distinct (Bridger et al., 2012; Voss & Federmeier, 2011). However, our results are agnostic to this debate. Our results focus on the facilitation in processing NP2 when NP1 shares phonological features with it, which is consistent with previous behavioral and ERP results (e.g., Desroches et al., 2009).

2.10.3. Retrieval results—A visual inspection of the grand average brain waves revealed two different patterns of brain responses based on age group. For younger adults, we observed a sustained negativity at the parietal sites between 700 ms to 1100 ms after pronoun onset. Importantly, this time window is consistent with the timing of the LPN effect reported in the recognition memory literature for retrieval of contextual (i.e., source) information, and continued evaluation of retrieval outcomes (Johansson & Mecklinger, 2003; Mecklinger et al., 2016). For older adults, however, we observed a *frontal* brain response emerging at approximately 400 ms after pronoun onset, and continuing until approximately the same time window as in younger adults. Based on this visual inspection and previous literature, we performed the statistical analyses in the 700–1100 ms time window, which was the common locus of effects for both groups (note the start of the effect is earlier for older adults, see General Discussion for a discussion on this temporal difference). Figure 2 shows grand average brain waves of younger (Panel A) and older adults (Panel B) during processing the critical pronouns for Phonologically Similar and Dissimilar conditions[9], and Table 2 reports the results of our analysis. As is clear from Table 2, we observed a significant interaction between Age and Phonological Similarity such that for younger adults, pronouns following phonologically similar NPs resulted in a greater negativity relative to pronouns following phonologically dissimilar NPs. However, for older adults, pronouns following phonologically similar NPs led to a smaller negativity compared to pronouns following phonologically dissimilar NPs (in Figure 2, compare the average of the two blue lines and the two red lines at the parietal regions in Panel A, and the at frontal regions in Panel B). We also observed a reliable interaction between Phonological Similarity and Brain Region such that pronouns following phonologically similar NPs resulted in smaller negativities at the frontal regions, but in greater negativity at the parietal regions. Finally, we observed a significant Age × Brain Region interaction such that the amplitudes of negativities were greater for younger than older adults at the frontal regions, but not at the central or parietal regions.

It is important to mention that we did not observe a main pronominal ambiguity effect, which has been previously reported for younger adults (Karimi et al., 2018; Nieuwland,

[9]Because Pronominal Ambiguity did not have an effect on averaged data (although it did have an effect in the expected direction for high-span individuals, see below), we are only showing the Phonological Similarity effect for simplicity. However, the plot for all 4 conditions is provided in the Supplementary Materials.

2014; Nieuwland & Van Berkum, 2006; van Berkum et al., 1999; Van Berkum et al., 2007). However, prior research has shown that this effect may vary as a function of working memory span (Nieuwland & Van Berkum, 2006). Therefore, we further analyzed the results by examining the pronominal ambiguity effect for high-span vs. low-span younger individuals. This analysis revealed a main Phonological Similarity × Pronominal Ambiguity × Brain Region interaction for high-span individuals ($F(2,34) = 4.48$ $p = .04$), but not for low-span individuals ($F(2,42) = .04$, $p = .87$). Follow-up simple effect analyses for high-span individuals revealed that ambiguous pronouns resulted in greater frontal negativities relative to unambiguous pronouns (at the 700–1100 ms time window), following phonologically dissimilar NPs ($F(1,17) = 8.12$, $p = .01$), but not following phonologically similar NPs ($F(1,17) = .26$, $p = .61$), which is consistent with previous reports (Nieuwland & Van Berkum, 2006). Because older adults had significantly smaller memory spans than younger adults ($t(76) = 2.45$, $p = .016$), the observation that older adults did not show a pronominal ambiguity effect is also in line with the finding that the pronominal ambiguity effect emerges only for high-span individuals (Nieuwland & Van Berkum, 2006). Moreover, the lack of an ambiguity effect for older adults is also consistent with previous research demonstrating that older adults are less efficient than younger adults when processing ambiguous pronouns during language comprehension (e.g., Dagerman, MacDonald, & Harm, 2006; Leonard, Waters & Caplan, 1997; Reifegerste, & Felser, 2017; Shake, & Stine-Morrow, 2011), perhaps due to "underspecified discourse model[s]" (Leonard, Waters & Caplan, 1997), or a general slow-down in information processing (Dagerman, MacDonald, & Harm, 2006). Interestingly, parallel results have been reported in language production, with older adults producing more ambiguous pronouns than younger adults, suggesting that they may "lack the necessary cognitive capacities to keep track of […] discourse referents" (Hendriks, Koster, & Hoeks, 2014).

2.10.4. Discussion of retrieval results—Younger and older adults showed different patterns of brain activity as a function of Phonological Similarity. Specifically, for younger adults, pronouns following similar-sounding NPs resulted in greater negativity relative to those following dissimilar-sounding NPs, suggesting that retrieval and/or integration difficulty varied as a function of phonological similarity. More specifically, this pattern of results implies that younger adults experienced more difficulty when retrieving and/or integrating information that is phonologically more confusable (i.e., similar-sounding), presumably due to similarity-based interference. These results provide support for theories assuming a phonologically mediated working memory (Acheson & MacDonald, 2011; Caramazza et al., 1981; Gibson, 1998; Shankweiler & Crain, 1986), and pose a challenge to the cue-based models of language processing (Kush et al., 2015; Lewis et al., 2006; Lewis & Vasishth, 2005). This is because cue-based retrieval models (in their current form) do not specify a mechanism for potential effects of phonological features on retrieval of stored memory representations (Kush et al., 2015; also see Jäger et al., 2017; Lewis et al., 2006; Lewis & Vasishth, 2005).

In contrast to the younger adults' results, for older adults, pronouns following phonologically similar NPs resulted in *smaller* negativities relative to pronouns following phonologically dissimilar NPs. Additionally, unlike the younger adults results, the timing and the topographical distribution of this effect resembles the Nref effect (van Berkum et al., 1999), which has more recently been shown to reflect retrieval difficulty of previously encoded representations during language processing (Barkley et al., 2015; Karimi et al., 2018). Because more negative Nref effects generally reflect *increased* difficulty, and this was observed following phonologically similar NPs, there are two potential explanations for this pattern of results. One possibility is that older adults actually experienced facilitated retrieval for pronouns following phonological similar NPs. A second possibility is that they resorted to some sort of good-enough/shallow processing whereby the referential dependency was not fully processed (i.e., both referential interpretations were not entertained). Given that we know older adults have more difficulty with memory retrieval, especially for phonological information (Burke et al., 1991, 2000; Diaz et al., 2014; Diaz et al., 2021; Taylor & Burke, 2002), and proper names (e.g., Cohen & Faulkner, 1986; Lovelace & Twohig, 1990; Ramscar et al., 2014), it is highly unlikely that the retrieval of the phonological and semantic representations of the referents were easier for older adults. Thus, we interpret our results as shallow processing on the part of the older adults when the preceding NPs were more phonologically similar. This shallow processing might mean they did not resolve the referential dependency at all, or predominantly resolved the pronoun in favor of one of the referential candidates to relieve the burden imposed on working memory by having to entertain two referential interpretations for two similar-sounding referential candidates. Note that because we did not observe an interaction between phonological similarity and pronominal ambiguity, this strategy seems to apply to both ambiguous and unambiguous pronouns.[10]

3. Post-hoc behavioral experiment: Methods

The results of the EEG experiment suggest that older adults might be performing shallow or good-enough processing whereby the referential dependency is either left unresolved or predominantly resolved to one candidate (without considering the alternative referential interpretation) when the two preceding referential candidates are phonologically similar (Nieuwland & Van Berkum, 2006). In order to directly test this possibility, we conducted a post-hoc behavioral experiment with the same design as the EEG experiment. Critically, the referent of the pronoun was probed after each trial to assess pronoun resolution accuracy for younger and older adults. If older adults underspecify the referential dependency, they should exhibit less accurate referential resolution when the two referential candidates sound alike relative to when they sound dissimilar.

3.1. Participants

Forty younger adults (Age range: 18–29, mean age = 19.66, 27 females, 13 males) and 40 older adults took part in the behavioral experiment, but we removed two older adults before the final analysis because their accuracy (in the Different Gender/Unambiguous conditions) fell below 60% (Age range: 60–78, mean age = 65.32, 24 females, 14 males). Older adults

[10]It is important to mention that we also investigated the potential influence of NP2 encoding (i.e., N400 amplitudes) on subsequent retrieval during pronoun processing (i.e., LPN amplitudes) for each participant. Although we did not observe any reliable correlations, given that these correlation analyses were based on only 40 data points (for each group), the results should not be taken as indicating a null relationship between encoding and retrieval. Future research with more participants is needed to further investigate this link.

were recruited online through Prolific (www.prolific.co). We restricted the older participants to monolingual native speakers of English residing in the United States, with "reliability" scores greater than 95%. They were paid $10 as compensation for their time (study duration = 28 – 58 minutes). Younger participants were recruited through the participant pool of Mississippi State University, and were compensated with course credit.

3.2. Stimuli, procedure and statistical analyses

We used the same critical stimuli and fillers as in the EEG experiment. The only difference was that each sentence was followed by a two-choice comprehension question. For critical sentences, the questions explicitly probed the referent of the pronoun, and the two referential candidates constituted the two response choices. For example, the question following *Hannah laughed at Jacob when he was almost drunk and high*, was: *Who was almost drunk and high?* The choices included the two preceding referential candidates (*Hannah* and *Jacob*). For fillers, the questions probed a specific part of the sentence. An example of a filler item would be: *The teacher hoped the students would stop talking while she was teaching*; Question: *What did the teacher hope the students would stop doing?;* Choices: *texting* - *talking*. The participants responded by clicking on one of the two response choices that appeared in two separate rows directly below the question. The positions of the two choices were randomized for each trial. The sentence presentation procedure was identical to that in the EEG experiment: Sentences were presented word by word, and each word stayed on the screen for 200 ms, followed by 300 ms of blank screen.

We ran mixed effects regression models on the accuracy and reaction times to the comprehension questions, using the (g)lmer function from the lme4 package in R (version 3.6.1). We always kept the random-effects structure maximal (Barr et al., 2013). That is, we always included random slopes for both subjects and items, as well as by-subjects and by-items random slopes for all the relevant predictors (note that by-subjects random slopes cannot be defined for the effect of Age). In case of non-convergence, we systematically simplified the random-effect structure by first removing the correlations between intercepts and slopes, and then removing the intercepts altogether (but keeping the random slopes; Barr et al., 2013). In any case, if two random effects were highly correlated, we dropped the one that captured less variance (Winter, 2020). For reaction time analyses, we first cleaned the data by removing any RTs faster than 1000 ms or slower than 10000 ms, and then further removing RTs that fell 2.5 standard deviations below or above the mean for each subject, which resulted loss of 556 out of 7800 data points (7.1%). To approach normality, we used the log-transformed RTs in the final models. Finally, the p-values for the reaction time analyses were calculated using the normal approximation method. The raw data and R script analysis scripts for all analyses can be found in the supplementary materials (https://osf.io/75qyz/).

3.3. Results

Because there are no accurate responses for the Same Gender/Ambiguous conditions (i.e., *she* or *he* could refer to either NP when both have the same gender), we performed separate analyses for the Same Gender/Ambiguous and the Different Genders/Unambiguous conditions: In the "Accuracy" analysis, we examined accuracy and reactions times for accurate responses in the two Different Gender/Unambiguous conditions, and in the "Interpretation Bias" analysis, we analyzed the tendency to interpret the ambiguous pronoun as referring to NP1 vs. NP2 and the associated reaction times in the Same Gender/ Ambiguous conditions. For brevity, we only report the statistically significant effects here, except for the critical effect of Phonological Similarity, which we report regardless of statistical significance.

3.3.1. Accuracy analysis—Our initial analysis showed no effect of Phonological Similarity on accuracy rates, a marginally significant effect of Age in the opposite direction (i.e., greater overall accuracy for older than younger adults), and no reliable interaction between Age and Phonological Similarity. However, given that the successful maintenance of the memory representations associated with the two referential candidates might may vary as function of the syntactic and/or linear positions of the candidates, we suspected older adults' accuracy might depend on whether the pronoun was co-indexed with the first- or the second-mentioned referential candidate. This is because past studies have demonstrated that first- and second-mentioned NPs confer different memory advantages. Specifically, first-mentioned NPs have been shown to be highly salient in memory because of their syntactically prominent role in English (e.g., Ariel, 2014; Arnold, 2001; Fukumura & Gompel, 2011; Fukumura & van Gompel, 2010; Givón, 1983; Gundel et al., 1993; Karimi, Diaz, & Ferreira, 2019; Karimi & Ferreira, 2016b; Karimi, Fukumura, Ferreira, & Pickering, 2014; Stevenson et al., 1994), and to enjoy increased attention and accessibility in short- as well as long-term memory (e.g., Birch, & Rayner, 2010; McKoon, Ratcliff, Ward, & Sproat, 1993). At the same time, past research has also shown that more recently encoded memory items are retrieved more easily at a subsequent point (i.e., the recency effect; e.g., Azizian & Polich, 2007; Davelaar, GoshenGottstein, Ashkenazi, Haarmann, & Usher, 2005; Glanzer & Cunitz, 1966; Murdock, 1962), perhaps because they are necessarily less susceptible to time-based decay (Lewis et al., 2006; Lewis & Vasishth, 2005). Thus, it is possible that although younger adults are able to maintain the representations of both referential candidates in memory, older adults may primarily have access to only one representation depending on whether syntactic prominence (NP1) or recency (NP2) has a stronger effect on the aging brain (in our experimental context). We thus investigated the effects of NP Position (NP1 vs. NP2) and Phonological Similarity for younger and older adults separately. Tables 3 and 4 report the accuracy rates for each condition as well as the statistical significance of the effects for younger and older adults, respectively.

For younger adults, there was no significant effects of NP Position ($\beta = -.00$, $SE = .31$, $z = .002$, $p = .99$), Phonological Similarity ($\beta = .33$, $SE = .30$, $z = 1.08$, $p = .28$), or their interaction ($\beta = -.91$, $SE = .58$, $z = -1.55$, $p = .12$). However, for older adults, although there were no main effects of NP Position ($\beta = .24$, $SE = .28$, $z = .86$, $p = .38$), and Phonological Similarity ($\beta = .24$, $SE = .27$, $z = .86$, $p = .38$), the interaction between these two variables was significant ($\beta = 1.19$, $SE = .55$, $z = 2.17$, $p = .02$). Follow-up analyses showed that while Phonological Similarity had no reliable effect on accuracy rates when the pronoun was co-indexed with the first-mentioned referential candidate (i.e., NP1; $\beta = -.70$, $SE = .54$, $z = -1.29$, $p = .19$), older adults were significantly less accurate in the Phonologically Similar condition relative to the Phonologically Dissimilar condition when the pronoun was co-indexed with the second-mentioned referential candidate (i.e., NP2; $\beta = 17.63$, $SE = 3.60$, $z = 4.89$, $p < .001$). Note that another way to view this interaction is that there was no effect of NP Position in Phonologically Similar conditions ($\beta = .64$, $SE = 1.47$, $z = .43$, $p = .66$), but older adults were more accurate when the pronoun referred to the second-mentioned (relative to when it referred to first-mentioned) referential candidate when the two NPs were Phonologically Dissimilar ($\beta = 26.44$, $SE = 4.07$, $z = 6.48$, $p < .001$).

We also analyzed RTs for accurate responses as a function of NP Position and Phonological Similarity for the two groups. Tables 5 and 6 report mean RTs for each condition for younger and older adults, respectively.

Mixed-effects regression analyses showed a main effect of Phonological Similarity for younger adults ($\beta = -.04$, $SE = .01$, $t = 2.37$, $p = .02$), with slower reactions times in the Phonologically Similar condition relative to the Phonologically Dissimilar condition. However, older adults showed no Phonological Similarity effect ($\beta = -.00$, $SE = .00$, $t = -.002$, $p = .99$).

3.3.2. Interpretation Bias—For the Gender Congruent/Ambiguous conditions, we analyzed the effects of Age and Phonological Similarity on the probability of interpreting the ambiguous pronoun as referring to NP1 or NP2. Table 7 reports the interpretation biases and the associated RTs for each NP for each condition.

Statistical analyses revealed no significant effects of Age or Phonological Similarity on NP1 vs. NP2 interpretations. However, we observed a significant main effect of phonological similarity on reaction times ($\beta = -.03$, $SE = .01$, $t = -2.92$, $p = .003$), with both groups being slower when the referential candidates were phonologically similar than when they were phonologically dissimilar.

3.4. Discussion

The post-hoc behavioral results showed that unlike younger adults, older adults' accuracy rates depended on whether the pronoun referred to the first-mentioned (NP1) or the second-mentioned referential candidate (NP2); while older adults' accuracy rates did not vary as a function of phonological similarity when the pronoun was co-indexed with the first-mentioned NP, their accuracy rates significantly dropped for pronouns following phonologically similar (relative to phonologically dissimilar) referential candidates when the pronoun was co-indexed with the second-mentioned NP. This pattern of results suggests that older adults' processing of the pronoun is shallow when the referential candidates sound similar, consistent with the good-enough processing theory. However, this good-enough processing takes place only when the pronoun is co-indexed with NP2 (and not with NP1). Note that older adults were also more accurate when the pronoun was co-indexed with NP2 compared to when it was co-indexed with NP1, when the two NP2s were phonologically dissimilar. Thus, taken together, accuracy rates for older adults were greater in the dissimilar than in the similar condition for NP2, as well as for NP2 than for NP1 in the dissimilar condition, which suggest that older adults have lopsided access to the two NPs, with a clear advantage for NP2 (the more recent referential candidate). Importantly, this pattern of results also speaks to the potential mechanism underlying the shallow processing for older adults;

they primarily maintain NP2 in memory, and may not retrieve NP1 as easily and/or in most trials, and as a result, although phonological similarity exerts an effect on NP2, it does not have an effect on NP1.

The reaction time results were also consistent with a good-enough account: For Different Genders/Unambiguous conditions, younger adults exhibited a Phonological Similarity effect (with slower reactions times for phonologically similar NPs), while older adults showed no such effect, suggesting that while younger adults are sensitive to phonological similarity, older adults are not, which is consistent with our EEG results. Coupled with the accuracy results, the lack of a phonological similarity effect for older adults' reaction times can be interpreted as them predominantly entertaining NP2 as the potential referent of the pronoun, thereby essentially bypassing the processing difficulty imposed by phonological similarity.

Finally, when the pronoun was ambiguous (Same Gender conditions), there was a main effect of phonological similarity with slower reaction times for pronouns following phonologically similar relative to those following phonologically dissimilar NPs, suggesting that phonological similarity caused processing difficulty for both groups. Note that this last result suggests that phonological similarity should result in processing difficulty (not ease), and therefore the lack of a phonological similarity effect for older adults' reaction times in the Different Genders/Unambiguous conditions, along with their lower accuracy rates for phonologically similar referential candidates when the pronoun was co-indexed with NP2, clearly suggest that older adults' referential processing was shallow during the EEG experiment. Also note that the post-hoc behavioral experiment illustrated shallow processing on the part of older adults even when the referential dependency was probed on *every* trial (which highlights the NPs). Since past research has shown that structural dependencies are significantly *more* likely to be resolved when questions directly assess them (Stewart, Holler, & Kidd, 2007; Swets et al., 2008), older adults may have been even more prone to good-enough processing during the EEG experiment because the questions never probed the referential dependency, and were asked infrequently.

Thus, importantly, the results of our post-hoc behavioral experiments not only provide direct evidence for shallow processing on the part of older adults, they also reveal a mechanism through which such shallow processing could occur. Specifically, our results suggest that the pronoun is probably not left unresolved by older adults when processing the sentences. If this was the case, we should have observed a main effect of Phonological Similarity, with significantly lower accuracy rates in the Similar relative to the Dissimilar condition. Rather, older adults seem to have a more robust representation of NP2, and therefore phonological similarity only had an effect when the felicitous referent was the second-mentioned NP. After all, if the representation of NP1 is not fully accessible (i.e., is degraded), phonological similarity is less likely to exert an influence on it. As mentioned above, a more robust representation for NP2 could be due to the greater accessibility of more recently encoded items in memory (i.e., NP2, e.g., Azizian & Polich, 2007; Davelaar, GoshenGottstein, Ashkenazi, Haarmann, & Usher, 2005; Glanzer & Cunitz, 1966; Murdock, 1962). Thus, based on our results, older adults are more susceptible to good-enough/shallow processing as a function of phonological similarity when the relevant memory items are more recently encoded in memory (such as in the case of second-mentioned NPs). That said, given that our study was not designed to uncover the mechanism underlying shallow processing on the part of older adults, future research is needed to further examine the mechanism through which older adults perform shallow processing.

One question that arises here is that if older adults have a more robust (less degraded) representation of NP2, why did this tendency only manifest itself in the Gender Incongruent/ Unambiguous conditions, but not in the Gender Congruent/Ambiguous conditions. One potential reason for this discrepancy is that when both referential candidates match the gender of the pronoun, participants are forced to entertain both referential interpretations to a greater degree compared to when only one referential candidate matches the gender of the pronoun, essentially removing the NP2 bias on the part of older adults.

4. General Discussion

In this experiment, we investigated the potential role of phonological similarity on the encoding and retrieval of referential candidates during sentence processing for younger and older adults. Critical sentences included two noun phrases in the form of proper names (NP1 and NP2) followed by a pronoun. The two names were either phonologically similar (*Jacob* and *Jason*/*Jade*) or dissimilar (*Jacob* and *Matt*/*Hannah*), and the pronoun either unambiguously referred to one of the NPs or was ambiguous between the two NPs (depending on their gender). By time-locking the ERPs to the onsets of NP2 and the pronoun, we were able to investigate the potential effect of phonological similarity on the encoding and retrieval of the memory representations of the referential candidates, respectively. In what follows, we will discuss the results of encoding and retrieval processes for younger and older adults.

4.1. Encoding phonologically similar words across younger and older adults

During encoding of the second NP, both younger and older adults exhibited a phonological priming effect whereby the encoding of the second NP (NP2) was easier when the preceding NP (NP1) was phonologically similar to it relative to the when NP1 was not phonologically similar. This pattern of results is consistent with the previous studies reporting easier lexical processing as a function of phonological overlap (e.g., Carreiras et al., 2005; Desroches et al., 2009; Humphreys et al., 1982; Lukatela & Turvey, 1994), and runs counter to the results reported by Kush et al. (2015). The underlying mechanism for phonological priming is argued to be that when two consecutive words share phonological features, processing the first one leads to the activation of its phonological features, and when the following word is encountered, the shared phonological features are already active, rendering processing easier. The fact that the phonological priming effect held for both younger and older adults reflects age-related stability in phonological priming. Although older adults often have increased difficulty with phonological retrieval, they are generally less impaired in basic aspects of reading, such as orthography-phonology conversions (Burke & Peters, 1986; Caplan & Waters, 2005; Stine & Wingfield, 1994; Verhaeghen, 2003). Since priming is supposed to be an automatic (as opposed to controlled) cognitive process, a broader implication of our results is that cognitive decline may minimally impact automatic cognitive processes such as priming. A relevant point here is the distance between the first and the second NP in our stimuli; because the maximum distance between the two NPs was four words (which amounts to 2000 ms temporally), our argument supporting maintenance of phonological priming during sentence processing is limited to fairly short temporal distances.

As mentioned above, the scalp distribution of this phonological priming effect was frontal, whereas the standard N400 component typically has a parietal distribution. However, the frontal distribution we observed is consistent with numerous previous studies that have reported frontal N400 effects as a function of conceptual and phonological priming, with primed stimuli producing smaller N400 amplitudes relative to unprimed stimuli (Jescheniak et al., 2002; Voss et al., 2010; Voss & Paller, 2006). There is an ongoing debate about whether frontal and parietal N400 components reflect distinct cognitive processes (memory recognition vs. semantic priming, respectively), with some researchers arguing for no functional distinction between the two (Voss & Federmeier, 2011), and some arguing for a different underlying processing for the two. However, our study was not designed to address this potential distinction.

The observation that shared phonological features facilitate the encoding of words is inconsistent with the results of Kush et al. (2015) who reported phonological encoding interference. As mentioned above, Kush et al. employed a paradigm in which participants had to maintain three words in memory while reading a sentence. Although reasonable for investigating the role of memory during sentence processing, such a paradigm imposes unnatural demands on the processor, and might change the nature of sentence comprehension itself. Normally, language processing does not involve keeping unrelated words in memory. Rather, it involves integrating the incoming words with the current sentence representation. Importantly, the additional words in Kush et al.'s study were irrelevant to the sentence meaning, which reduces the likelihood that participants were integrating them with the evolving sentence representation. In contrast, our design required that NP2s be integrated with the current sentence representation for successful comprehension. The task demands in Kush et al.'s memory load paradigm may unduly emphasize memory encoding, and may have presented participants with a dual-task, limiting the extent to which the incoming words were *integrated* into the discourse representation. Another reason for the inconsistency between Kush et al.'s and our results is that they employed rhyming stimuli whereas we used both onset- as well as rhyme-overlapping NPs. Critically, although our effects were in the same direction for both overlap types, they were slightly smaller for rhyming words, suggesting that rhyming words might induce smaller perceived similarity than onset-overlapping words (Fricke et al., 2016). Thus, the apparent discrepancy between our results and those of Kush et al. (2015) are likely due to the emphasis on memory encoding in their design and on integration in ours.

4.2. Retrieving phonologically similar words across younger and older adults

Unlike encoding, younger and older adults exhibited different patterns of brain activity during the retrieval of phonologically similar relative to phonologically dissimilar words. Specifically, for younger adults, pronouns following two phonologically similar NPs resulted in greater negativities compared to pronouns following two phonologically dissimilar NPs in the time window spanning from 700 to 1100 ms after the onset of the critical pronoun. However, older adults exhibited a reverse Nref effect emerging at about 400 ms after pronoun onset (i.e., the retrieval cue), whereby pronouns following two phonologically similar NPs produced *smaller* negativities relative to pronouns following two phonologically dissimilar NPs. This clearly demonstrates that younger adults experienced difficulty retrieving and/or integrating the referent when the referential candidates were phonologically similar, whereas older adults exhibited an apparent facilitation under the same scenario.

We outlined two potential accounts for the effect of phonology during sentence processing based on previous research findings. Under one account, working memory is phonologically-mediated and therefore phonological similarity should increase the retrieval difficulty of associated words (Caramazza et al., 1981; Gibson, 1998; Shankweiler & Crain, 1986). Alternatively, based on the cue-based models of language processing, previously encoded items are primarily accessed via semantic/morpho-syntactic, but not phonological cues, and therefore phonological similarity should have no effect on later retrieval (Lewis et al., 2006; Lewis & Vasishth, 2005).

Consistent with the phonologically mediated view of working memory (e.g., Acheson & MacDonald, 2011), and with previous research from the general memory literature (Mecklinger et al., 2007, 2016), our results showed that, younger adults experienced difficulty when processing the pronouns, with greater amplitudes for pronouns following phonologically similar NPs compared to those following phonologically dissimilar NPs. Importantly, this pattern of results is consistent with previous research from the memory literature showing that successful recall of source/contextual information, and continued evaluations of retrieval outputs also results in greater negativities in about the same time window and in parietal scalp regions (e.g., Cruse & Wilding, 2009; Johansson & Mecklinger, 2003; Mecklinger et al., 2007, 2016; Wilding, 2000). Moreover, previous research has demonstrated recruitment of general recognition memory processes for successful retrieval of referential candidates during pronoun processing, localized, at least partially, to the posterior parietal regions (Nieuwland et al., 2007; Nieuwland & Martin, 2017). Thus, to the extent that phonological similarity complicates subsequent recognition, phonological similarity should elicit similar brain responses as successful recognition of previously encoded items. Because the timing, direction, and the topographical distribution of our phonological effects matches with reported effects in the general memory literature, and because there is empirical evidence supporting the engagement of general recognition memory systems during referential processing per se (Nieuwland et al., 2007; Nieuwland & Martin, 2017), our results could be interpreted as showing that retrieving phonological information engages similar memory processes as recognition memory in general, and source/contextual retrieval and/or evaluation of retrieval outputs in particular. Retrieving phonological and source information might engage the same processes because both types of information are not central to a memory representation. Specifically, contextual memory involves retrieving attributes of the context rather than main item studied. Likewise, phonological information includes details about words’ sound forms rather than their meaning which is more central to communication than word forms.

Somewhat surprisingly, older adults exhibited the opposite pattern of results, namely, greater frontal negativities for pronouns following phonologically dissimilar relative to phonologically similar words. As mentioned above, such sustained, frontal negativities emerging at about 400 ms following a retrieval cue have been shown to reflect the retrieval difficulty of previously-encoded representations (i.e., an Nref effect, (Barkley et al., 2015; Karimi et al., 2018). Thus, this pattern of results for older adults has two possible interpretations: Retrieval under phonological similarity (i.e., more confusability) was easier for older adults, or older adults engaged in shallower processing to save computational resources when the referential candidates sounded similar. Given that older adults have more difficulty with phonological retrieval (not less, Burke et al., 1991, 2000; Diaz et al., 2014; 2021), and retrieving proper names (Cohen & Faulkner, 1986; Lovelace & Twohig, 1990; Ramscar et al., 2014), we argued that the more likely interpretation of the older adults' retrieval results is shallow processing. Such an account is in line with the Good Enough approach to language processing according to which processing resources are only allocated to the extent that is required by the task at hand (Ferreira, 2003; Ferreira et al., 2002; Karimi & Ferreira, 2016a;Sanford & Sturt, 2002; Swets et al., 2008).

4.3. Direct evidence for good-enough/shallow processing on the part of older adults

Our post-hoc behavioral results provided direct evidence that older adults were performing good-enough or shallow processing when resolving the referential dependency. There were three pieces of evidence leading to this interpretation: First, for ambiguous pronouns, both younger and older adults exhibited processing difficulty (as reflected by longer reaction times) when the two referential candidates sounded alike compared to when they sounded different, suggesting that phonological similarity increases processing difficulty for both groups. Second, for unambiguous pronouns, only younger adults (but not older adults) exhibited a phonological similarity effect where phonologically similar NPs produced greater processing difficulty (as reflected by longer reaction times), suggesting that younger adults are sensitive to phonological similarity but older adults are not. Third, when unambiguous pronouns referred to the second-mentioned (more recently encoded) NPs, older adults were significantly less accurate when the referential candidates sounded alike than when they sounded dissimilar, suggesting that older adults do not successfully resolve pronouns referring to the second-mentioned referential candidates. If phonological similarity produces processing difficulty, as suggested by the pattern of results in the ambiguous conditions, older adults should have shown the same phonological similarity effect for the unambiguous pronouns too. However, older adults showed both less accuracy and no sensitivity to phonological similarity in the different genders/unambiguous conditions, suggesting that they do not fully process the pronoun when the preceding referential candidates sound alike.

Two potential mechanisms could have led to good-enough pronominal processing for older adults: First, it is possible that they left the referential dependency unresolved. This means that they did not retrieve any representations when the NPs sounded similar, but retrieved the representation associated with the referent when they sounded different, leading to greater frontal negativities when the referential candidates were phonologically dissimilar compared to when they were phonologically similar. Another possibility is that when the two candidates were phonologically similar, older adults quickly resolved the pronoun to one of the referential candidates, thereby circumventing the need to entertain two referential interpretations in memory. In addition to providing direct evidence for shallow processing on the part of older adults, the results of our post-hoc behavioral experiments also provided some evidence for the mechanism through which shallow processing might have occurred in our experiments. Specifically, our results suggest that rather than leaving the pronoun unresolved, older adults seem to primarily entertain the second-mentioned NP as the referent of the pronoun, perhaps because they have a more robust (less degraded) memory representation for the second-mentioned (more recently encoded) NP. Consequently, the effect of phonological similarity only applies to NP2, leading to lower accuracy rates when the pronoun was co-indexed with NP2. In other words, older adults essentially bypass the processing burden imposed by similar-sounding words by primarily entertaining NP2 as the referent of the pronoun. Consistent with these results, phonological similarity has been shown to disrupt memory for the *order* of encoded items (as opposed to the items themselves; Conrad & Hull, 1964; Fallon et al., 1999; Wickelgren, 1965). Interestingly, because serial order has been shown to have a considerable influence on memory activation and therefore pronoun processing (Ariel, 2014; Arnold, 2001; Fukumura & Gompel, 2011; Fukumura & van Gompel, 2010; Givón, 1983; Gundel et al., 1993; Stevenson et al., 1994), our results suggest that older adults were more strongly affected by the disruption of serial order memory, leading to shallow processing.

4.4. Age-dependent differential effects of phonological similarity at retrieval

An important question that might arise at this point is why younger and older adults exhibited different patterns of brain activity when processing the pronoun (the retrieval cue). The different processes involved in referential processing might help illuminate this issue. Based on a prominent model of referential processing proposed by Garrod and colleagues (Garrod & Sanford, 1990; 1994; Garrod & Terras, 2000; Sanford, Garrod, Lucas, & Henderson, 1983; also see Coopmans & Nieuwland, 2020), referential processing involves two distinct stages: An initial “bonding” stage where the representations associated with potential referents are activated, and a tentative and superficial, referential bond is formed between the referring expression (the pronoun, in our case) and one of the potential referents, and a following “resolution” stage where the output of the bonding stage is evaluated against the context and the general discourse model. Recently, a two-stage mechanism for referential processing involving an initial activation and a subsequent integration of referents has been corroborated by demonstrating distinct brain oscillations for these two processes (Coopmans & Nieuwland, 2020). Based on our results, older adults’ brain responses seem to reflect difficulty with the “bonding” stage, and younger adults’ brain responses seem to reflect difficulty at the “resolution” stage. Specifically, the reactivation/retrieval[11] of the referential candidates might be difficult for older adults, which in turn seems to lead to shallow processing. However, while younger adults seem to have no difficulty with reactivation/retrieval of referential candidates, they seem to

[11]Note that we are assuming no distinction between reactivation and retrieval. The most highly activated representation is automatically the one that is “retrieved”. In other words, reactivation and retrieval are not two separate operations. There is only reactivation, and the representation with the greatest activation level is automatically linked to the pronoun.

experience difficulty when integrating the activated/retrieved representation when there is a phonologically similarity competitor. This interpretation of the younger adults' data is fully consistent with findings from the memory literature showing the LPN reflects the continued evaluation of retrieval outputs (e.g., Johansson & Mecklinger, 2003; Mecklinger et al., 2016). This interpretation is also fully consistent with the timing of the effects. Specifically, the reverse Nref effect for older adults emerged earlier than the LPN effect for younger adults (see Figure 2). Given that bonding is supposed to precede resolution, the earlier-emerging Nref effect exhibited by older adults, and the later LPN effect exhibited by younger adult map onto bonding and resolution stages, respectively. Note that we proposed two potential mechanisms for how shallow referential processing might take place for older adults (i.e., non-resolution of the referential dependency, or resolution to one referent on a random basis, see above). To the extent that older adults only re-activate the potential referents and bypass the integration stage, they are not actually resolving the pronoun. Thus, in this sense, a bonding-only mechanism would be in line with non-resolution of the pronoun rather than resolution on a random basis.

4.5. Implications for cue-based retrieval theories

Our results have important implications for cue-based models of language processing. According to these models, incoming information rapidly fades from working memory and is stored in long-term memory, access to which is made possible through semantic and/or morpho-syntactic, but not phonological information. As a result, based on this view, although phonological information might affect the encoding of incoming words, it should not affect retrieval (Kush et al., 2015). However, our results clearly showed that phonological information does affect retrieval, such that phonological similarity between to-be-retrieved memory items renders retrieval more difficult for younger adults. Specifically, our results suggest that when the morpho-syntactic features of the critical pronoun re-activated the potential referential candidates, the phonological features of those referential candidates were also reactivated, and caused difficulty with the retrieval of the target referential candidate (i.e., the referent) when the activated phonological features were similar. Thus, our results suggest that although phonological features might not serve as retrieval cues, the retrieval process itself is affected by phonological information such that overlapping phonological features render the retrieval operation more difficult. Thus, although our results by no means call into question the overall validity of cue-based models, they do suggest that these models could be revised to incorporate phonological effects, as well as age-induced shallow processing. Another implication of our results for the cue-based models concerns syntactic/linear position. Our results suggest that the probability of age-induced shallow processing is greater for more recently mentioned memory items. One way to interpret these results is that more recently encoded memory items are less degraded in memory and are therefore also more prone to phonological interference effects. To the extent that this is true, cue-based models might incorporate mechanisms for differential interference effects as a function of differential decay rates caused by cognitive decline and syntactic/linear position.

4.6. Consistency with previous findings and limitations

We also replicated previous findings showing an effect of pronominal ambiguity for high reading span young individuals (Nieuwland & Van Berkum, 2006). Specifically, high-span (but not low-span) younger adults exhibited a pronominal ambiguity effect with greater frontal and sustained negativity for ambiguous relative to unambiguous pronouns when the NPs were phonologically dissimilar. This effect has been attributed to having to entertain two referential interpretations in the case of ambiguous pronouns vs. one referential interpretation in the case of unambiguous pronouns. Specifically, it has been argued that when processing ambiguous pronouns such as the ones in our stimuli, high-span individuals maintain both referential interpretations, whereas low-span individuals quickly resolve the pronoun to the more accessible candidate, presumably to save computational resources, (Nieuwland & Van Berkum, 2006).

One limitation of the current study is that the generalizability of the results is limited to proper names. However, the results do show the importance of investigating phonological factors during language processing, as such a factor would offer valuable insight to our understanding of the human sentence processing machinery. Moreover, our results highlight the importance of examining psycholinguistic effects across the life span to obtain a better understanding of whether and how the human language processing system changes with age.

5. Conclusions

We investigated the encoding and retrieval of phonologically similar words for younger and older adults during sentence processing. The results showed that phonological similarity leads to priming for both younger and older adults, and therefore facilitated integration during encoding, which demonstrates consistency in phonological priming across the lifespan. However, during retrieval, younger adults exhibited retrieval difficulty as a function of phonological similarity whereas older adults may have left the ambiguity unresolved to relieve memory load, consistent which the Good Enough approach to language processing (Ferreira et al., 2002). The results clearly show that phonological information affects both encoding and retrieval during sentence processing and challenge current cue-based retrieval theories.

Acknowledgements

This project was funded by NIH Grant R01 AG034138 (M.T.D.). We thank Caitlin Little, Emma Domico, and Peter Weber for assistance with data collection and coding. We also thank the staff and scientists at the Center for Language Science (CLS), the Social, Life, and Engineering Sciences Imaging Center (SLEIC), and the Language and Aging lab at the Pennsylvania State University for their support and feedback. The authors declare no conflicts of interest, financial or otherwise, that would preclude a fair review or publication of this manuscript. The full list of stimuli, the raw data, as well as the full analysis results are uploaded to the Open Science Framework, and can be accessed at: https://osf.io/75qyz/.

References

Abrams L, & Rodriguez EL (2005). Syntactic class influences phonological priming of tip-of-the-tongue resolution. Psychonomic Bulletin & Review, 12(6), 1018–1023. [PubMed: 16615322]

Acheson DJ, & MacDonald MC (2011). The rhymes that the reader perused confused the meaning: Phonological effects during on-line sentence comprehension. Journal of Memory and Language, 65(2), 193–207. 10.1016/j.jml.2011.04.006 [PubMed: 21743771]

Ariel M (2014). Accessing Noun-Phrase Antecedents (RLE Linguistics B: Grammar). Routledge. 10.4324/9781315857473

Arnold JE (2001). The Effect of Thematic Roles on Pronoun Use and Frequency of Reference Continuation. Discourse Processes, 31(2), 137–162. 10.1207/S15326950DP3102_02

Azizian A, & Polich J (2007). Evidence for attentional gradient in the serial position memory curve from event-related potentials. Journal of Cognitive Neuroscience, 19(12), 2071–2081. [PubMed: 17892393]

Baddeley AD (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. Quarterly Journal of Experimental Psychology, 18(4), 302–309. 10.1080/14640746608400047

Baddeley AD (2018). Short-term Memory for Word Sequences as a Function of Acoustic, Semantic and Formal Similarity: Quarterly Journal of Experimental Psychology. 10.1080/14640746608400055

Barkley C, Kluender R, & Kutas M (2015). Referential processing in the human brain: An Event-Related Potential (ERP) study. Brain Research, 1629, 143–159. 10.1016/j.brainres.2015.09.017 [PubMed: 26456801]

Barr DJ, Levy R, Scheepers C, & Tily HJ (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278. 10.1016/j.jml.2012.11.001

Bowles NL, & Poon LW (1985). Aging and Retrieval of Words in Semantic Memory. Journal of Gerontology, 40(1), 71–77. 10.1093/geronj/40.1.71 [PubMed: 3965563]

Bridger EK, Bader R, Kriukova O, Unger K, & Mecklinger A (2012). The FN400 is functionally distinct from the N400. NeuroImage, 63(3), 1334–1342. 10.1016/j.neuroimage.2012.07.047 [PubMed: 22850570]

Brown JI (1960). The Nelson-Denny Reading Test. Houghton Mifflin.

Brown R, & McNeill D (1966). The "tip of the tongue" phenomenon. Journal of Verbal Learning and Verbal Behavior, 5(4), 325–337. 10.1016/S0022-5371(66)80040-3

Burke DM, MacKay DG, & James LE (2000). Theoretical approaches to language and aging. In Models of cognitive aging (pp. 204–237). Oxford University Press.

Burke DM, MacKay DG, Worthley JS, & Wade E (1991). On the tip of the tongue: What causes word finding failures in young and older adults? Journal of Memory and Language, 30(5), 542–579. 10.1016/0749-596X(91)90026-G

Burke DM, & Peters L (1986). Word associations in old age: Evidence for consistency in semantic encoding during adulthood. Psychology and Aging, 1(4), 283–292. 10.1037/0882-7974.1.4.283 [PubMed: 3267408]

Caplan D, & Waters G (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. Memory, 13(3–4), 403–413. 10.1080/09658210344000459 [PubMed: 15952262]

Caramazza A, Berndt RS, Basili annamaria G., & Koller JJ (1981). Syntactic Processing Deficits in Aphasia. Cortex, 17(3), 333–347. 10.1016/S0010-9452(81)80021-4 [PubMed: 7333108]

Carreiras M, Ferrand L, Grainger J, & Perea M (2005). Sequential Effects of Phonological Priming in Visual Word Recognition. Psychological Science, 16(8), 585–589. 10.1111/j.1467-9280.2005.01579. [PubMed: 16102059]

Carreiras M, Gernsbacher MA, & Villa V (1995). The advantage of first mention in Spanish. Psychonomic Bulletin & Review, 2(1), 124–129. [PubMed: 24203596]

Conrad R, & Hull AJ (1964). Information, Acoustic Confusion and Memory Span. British Journal of Psychology, 55(4), 429–432. 10.1111/j.2044-8295.1964.tb00928.x [PubMed: 14237884]

Craik FIM (1968). Two components in free recall. Journal of Verbal Learning and Verbal Behavior, 7(6), 996–1004. 10.1016/S0022-5371(68)80058-1

Craik Fergus I. M., & Rose NS (2012). Memory encoding and aging: A neurocognitive perspective. Neuroscience & Biobehavioral Reviews, 36(7), 1729–1739. 10.1016/j.neubiorev.2011.11.007 [PubMed: 22155274]

Cross ES, & Burke DM (2004). Do alternative names block young and older adults' retrieval of proper names? Brain and Language, 89(1), 174–181. 10.1016/S0093-934X(03)00363-8 [PubMed: 15010248]

Cruse D, & Wilding EL (2009). Prefrontal cortex contributions to episodic retrieval monitoring and evaluation. Neuropsychologia, 47(13), 2779–2789. 10.1016/j.neuropsychologia.2009.06.003 [PubMed: 19523968]

Daneman M, & Carpenter PA (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19(4), 450–466. 10.1016/S0022-5371(80)90312-6

Davelaar EJ, Goshen-Gottstein Y, Ashkenazi A, Haarmann H, & Usher M (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. Psychological Review, 117(1), 3–42.

Dell GS, McKoon G, & Ratcliff R (1983). The activation of antecedent information during the processing of anaphoric reference in reading. Journal of Verbal Learning and Verbal Behavior, 22(1), 121–132. 10.1016/S0022-5371(83)80010-3

Desroches AS, Newman RL, & Joanisse MF (2009). Investigating the Time Course of Spoken Word Recognition: Electrophysiological Evidence for the Influences of Phonological Similarity. Journal of Cognitive Neuroscience, 21(10), 1893–1906. 10.1162/jocn.2008.21142 [PubMed: 18855555]

Diaz MT, Johnson MA, Burke DM, & Madden DJ (2014). Age-related Differences in the Neural Bases of Phonological and Semantic Processes. Journal of Cognitive Neuroscience, 26(12), 2798–2811. 10.1162/jocn_a_00665 [PubMed: 24893737]

Diaz MT, Johnson MA, Burke DM, Truong T-K, & Madden DJ (2019). Age-related differences in the neural bases of phonological and semantic processes in the context of task-irrelevant information. Cognitive, Affective, & Behavioral Neuroscience, 19(4), 829–844. 10.3758/s13415-018-00671-2

Diaz MT, Karimi H, Troutman SB, Gertel VH, Cosgrove AL, & Zhang H (2021). Neural sensitivity to phonological characteristics is stable across the lifespan. NeuroImage, 225, 117511. [PubMed: 33129928]

Fallon AB, Groves K, & Tehan G (1999). Phonological Similarity and Trace Degradation in the Serial Recall Task: When CAT helps RAT, but not MAN. International Journal of Psychology, 34(5–6), 301–307. 10.1080/002075999399602

Ferreira F (2003). The misinterpretation of noncanonical sentences. Cognitive Psychology, 47(2), 164–203. 10.1016/S0010-0285(03)00005-7 [PubMed: 12948517]

Ferreira F, Bailey KGD, & Ferraro V (2002). Good-Enough Representations in Language Comprehension. Current Directions in Psychological Science, 11(1), 11–15. JSTOR.

Fodor JD (1978). Parsing Strategies and Constraints on Transformations. Linguistic Inquiry, 9(3), 427–473. JSTOR.

Folstein MF, Folstein SE, & McHugh PR (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12(3), 189–198. [PubMed: 1202204]

Fricke M, Baese-Berk MM, & Goldrick M (2016). Dimensions of similarity in the mental lexicon. Language, Cognition and Neuroscience, 31(5), 639–645. 10.1080/23273798.2015.1130234

Fukumura K, & Gompel RPG van. (2011). The effect of animacy on the choice of referring expression. Language and Cognitive Processes, 26(10), 1472–1504. 10.1080/01690965.2010.506444

Fukumura K, & van Gompel RPG (2010). Choosing anaphoric expressions: Do people take into account likelihood of reference? Journal of Memory and Language, 62(1), 52–66. 10.1016/j.jml.2009.09.001

Gernsbacher MA (1989). Mechanisms that improve referential access. Cognition, 32(2), 99–156. 10.1016/0010-0277(89)90001-2 [PubMed: 2752708]

Gernsbacher MA, & Hargreaves DJ (1988). Accessing sentence participants: The advantage of first mention. Journal of memory and language, 27(6), 699–717. [PubMed: 25520541]

Gerrig RJ, & McKoon G (1998). The readiness is all: The functionality of memory- based text processing. Discourse Processes, 26(2–3), 67–86. 10.1080/01638539809545039

Gibson E (1998). Linguistic complexity: Locality of syntactic dependencies. Cognition, 68(1), 1–76. 10.1016/S0010-0277(98)00034-1 [PubMed: 9775516]

Givón T (1983). Topic Continuity in Discourse: A quantitative cross-language study. John Benjamins Publishing.

Glanzer M, & Cunitz AR (1966). Two storage mechanisms in free recall. Journal of verbal learning and verbal behavior, 5(4), 351–360.

Gordon PC, Hendrick R, & Johnson M (2001). Memory interference during language processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 27(6), 1411–1423. 10.1037//0278-7393.27.6.1411

Gordon PC, Hendrick R, Johnson M, & Lee Y (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(6), 1304–1321. 10.1037/0278-7393.32.6.1304

Gordon PC, Hendrick R, & Levine WH (2002). Memory-Load Interference in Syntactic Processing. PSYCHOLOGICAL SCIENCE, 13(5), 6.

Gundel JK, Hedberg N, & Zacharski R (1993). Cognitive Status and the Form of Referring Expressions in Discourse. Language, 69(2), 274–307. JSTOR. 10.2307/416535

Hofmeister P (2011). Representational complexity and memory retrieval in language comprehension. Language and Cognitive Processes, 26(3), 376–405. 10.1080/01690965.2010.492642 [PubMed: 22675235]

Humphreys GW, Evett LJ, & Taylor DE (1982). Automatic phonological priming in visual word recognition. Memory & Cognition, 10(6), 576–590. 10.3758/BF03202440 [PubMed: 7162419]

Jäger LA, Engelmann F, & Vasishth S (2015). Retrieval interference in reflexive processing: Experimental evidence from Mandarin, and computational modeling. Frontiers in Psychology, 6. 10.3389/fpsyg.2015.00617

Jäger LA, Engelmann F, & Vasishth S (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. Journal of Memory and Language, 94, 316–339. 10.1016/j.jml.2017.01.004

Jescheniak JD, Schriefers H, Garrett MF, & Friederici AD (2002). Exploring the Activation of Semantic and Phonological Codes during Speech Planning with Event-Related Brain Potentials. Journal of Cognitive Neuroscience, 14(6), 951–964. 10.1162/089892902760191162 [PubMed: 12191461]

Johansson M, & Mecklinger A (2003). The late posterior negativity in ERP studies of episodic memory: Action monitoring and retrieval of attribute conjunctions. Biological Psychology, 64(1), 91–117. 10.1016/S0301-0511(03)00104-2 [PubMed: 14602357]

Karimi H, Diaz M, & Ferreira F (2019). “A cruel king” is not the same as “a king who is cruel”: Modifier position affects how words are encoded and retrieved from memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 45(11), 2010.

Karimi H, & Ferreira F (2016a). Good-enough linguistic representations and online cognitive equilibrium in language processing. Quarterly Journal of Experimental Psychology, 69(5), 1013–1040. 10.1080/17470218.2015.1053951

Karimi H, & Ferreira F (2016b). Informativity renders a referent more accessible: Evidence from eyetracking. Psychonomic bulletin & review, 23(2), 507–525. [PubMed: 26260676]

Karimi H, Fukumura K, Ferreira F, & Pickering MJ (2014). The effect of noun phrase length on the form of referring expressions. Memory & cognition, 42(6), 993–1009. [PubMed: 24596114]

Karimi H, Swaab TY, & Ferreira F (2018). Electrophysiological evidence for an independent effect of memory retrieval on referential processing. Journal of Memory and Language, 102, 68–82. 10.1016/j.jml.2018.05.003

Kemper S, Kynette D, & Norman S (1992). Age Differences in Spoken Language. In West RL & Sinnott JD (Eds.), Everyday Memory and Aging: Current Research and Methodology (pp. 138–152). Springer. 10.1007/978-1-4613-9151-7_9

Kluender R, & Kutas M (1993). Bridging the Gap: Evidence from ERPs on the Processing of Unbounded Dependencies. Journal of Cognitive Neuroscience, 5(2), 196–214. 10.1162/jocn.1993.5.2.196 [PubMed: 23972154]

Koornneef A, & Vanberkum J (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. Journal of Memory and Language, 54(4), 445–465. 10.1016/j.jml.2005.12.003

Kush D, Johns CL, & Van Dyke JA (2015). Identifying the role of phonology in sentence-level reading. Journal of Memory and Language, 79–80, 18–29. 10.1016/j.jml.2014.11.001

Kutas M, & Federmeier KD (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). Annual Review of Psychology, 62(1), 621–647. 10.1146/annurev.psych.093008.131123

Kutas M, & Hillyard SA (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. Biological Psychology, 11(2), 99–116. 10.1016/0301-0511(80)90046-0 [PubMed: 7272388]

Lewis RL, & Vasishth S (2005). An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. Cognitive Science, 29(3), 375–419. 10.1207/s15516709cog0000_25 [PubMed: 21702779]

Lewis RL, Vasishth S, & Van Dyke JA (2006). Computational principles of working memory in sentence comprehension. Trends in Cognitive Sciences, 10(10), 447–454. 10.1016/j.tics.2006.08.007 [PubMed: 16949330]

Lucas MM, Tanenhaus MK, & Carlson GN (1990). Levels of representation in the interpretation of anaphoric reference and instrument inference. Memory & Cognition, 18(6), 611–631. 10.3758/BF03197104 [PubMed: 2266863]

Lukatela G, & Turvey MT (1994). Visual lexical access is initially phonological: 2. Evidence from phonological priming by homophones and pseudohomophones. Journal of Experimental Psychology: General, 123(4), 331–353. 10.1037/0096-3445.123.4.331 [PubMed: 7996120]

MacDonald MC, & MacWhinney B (1990). Measuring inhibition and facilitation from pronouns. Journal of Memory and Language, 29(4), 469–492. 10.1016/0749-596X(90)90067-A

MacKay DG, & James LE (2004). Sequencing, Speech Production, and Selective Effects of Aging on Phonological and Morphological Speech Errors. Psychology and Aging, 19(1), 93–107. 10.1037/0882-7974.19.1.93 [PubMed: 15065934]

McElree B (2000). Sentence Comprehension Is Mediated by Content-Addressable Memory Structures. Journal of Psycholinguistic Research, 29(2), 111–123. 10.1023/A:1005184709695 [PubMed: 10709178]

McElree B, Foraker S, & Dyer L (2003). Memory structures that subserve sentence comprehension. Journal of Memory and Language, 48(1), 67–91. 10.1016/S0749-596X(02)00515-6

Mecklinger A, Johansson M, Parra M, & Hanslmayr S (2007). Source-retrieval requirements influence late ERP and EEG memory effects. Brain Research, 1172, 110–123. 10.1016/j.brainres.2007.07.070 [PubMed: 17822684]

Mecklinger A, Rosburg T, & Johansson M (2016). Reconstructing the past: The late posterior negativity (LPN) in episodic memory studies. Neuroscience & Biobehavioral Reviews, 68, 621–638. 10.1016/j.neubiorev.2016.06.024 [PubMed: 27365154]

Murdock BB Jr (1962). The serial position effect of free recall. Journal of experimental psychology, 64(5), 482.

Nicol J, & Swinney D (1989). The role of structure in coreference assignment during sentence comprehension. Journal of Psycholinguistic Research, 18(1), 5–19. 10.1007/BF01069043 [PubMed: 2647962]

Nieuwland MS (2014). “Who’s he?” Event-related brain potentials and unbound pronouns. Journal of Memory and Language, 76, 1–28. 10.1016/j.jml.2014.06.002

Nieuwland MS, & Martin AE (2017). Neural Oscillations and a Nascent Corticohippocampal Theory of Reference. Journal of Cognitive Neuroscience, 29(5), 896–910. 10.1162/jocn_a_01091 [PubMed: 28129065]

Nieuwland MS, Petersson KM, & Van Berkum JJA (2007). On sense and reference: Examining the functional neuroanatomy of referential processing. NeuroImage, 37(3), 993–1004. 10.1016/j.neuroimage.2007.05.048 [PubMed: 17611124]

Nieuwland MS, & Van Berkum JJA (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. Brain Research, 1118(1), 155–167. 10.1016/j.brainres.2006.08.022 [PubMed: 16956594]

Osterhout L, & Swinney DA (1993). On the temporal course of gap-filling during comprehension of verbal passives. Journal of Psycholinguistic Research, 22(2), 273–286. 10.1007/BF01067834 [PubMed: 8366478]

Patil U, Vasishth S, & Lewis RL (2016). Retrieval Interference in Syntactic Processing: The Case of Reflexive Binding in English. Frontiers in Psychology, 7. 10.3389/fpsyg.2016.00329

Reuter-Lorenz PA, & Cappell KA (2008). Neurocognitive Aging and the Compensation Hypothesis. Current Directions in Psychological Science, 17(3), 177–182. 10.1111/j.1467-8721.2008.00570.x

Rizio AA, Moyer KJ, & Diaz MT (2017). Neural evidence for phonologically based language production deficits in older adults: An fMRI investigation of age-related differences in picture-word interference. Brain and Behavior, 7(4), e00660. 10.1002/brb3.660 [PubMed: 28413708]

Sanford AJ, & Garrod SC (1989). What, when, and how?: Questions of immediacy in anaphoric reference resolution. Language and Cognitive Processes, 4(3–4), SI235–SI262. 10.1080/01690968908406369

Sanford Anthony J., & Garrod SC (2005). Memory-Based Approaches and Beyond. Discourse Processes, 39(2–3), 205–224. 10.1080/0163853X.2005.9651680

Sanford Anthony J., & Sturt P (2002). Depth of processing in language comprehension: Not noticing the evidence. Trends in Cognitive Sciences, 6(9), 382–386. 10.1016/S1364-6613(02)01958-7 [PubMed: 12200180]

Shankweiler D, & Crain S (1986). Language mechanisms and reading disorder: A modular approach. Cognition, 24(1), 139–168. 10.1016/0010-0277(86)90008-9 [PubMed: 3791919]

Stanovich KE, & West RF (1989). Exposure to Print and Orthographic Processing. Reading Research Quarterly, 24(4), 402–433. JSTOR. 10.2307/747605

Stevenson RJ, Crawley RA, & Kleinman D (1994). Thematic roles, focus and the representation of events. Language and Cognitive Processes, 9(4), 519–548. 10.1080/01690969408402130

Stine EAL, & Wingfield A (1994). Older adults can inhibit high-probability competitors in speech recognition. Aging, Neuropsychology, and Cognition, 1(2), 152–157. 10.1080/09289919408251456

Swets B, Desmet T, Clifton C, & Ferreira F (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. Memory & Cognition, 36(1), 201–216. 10.3758/MC.36.1.201 [PubMed: 18323075]

Taylor JK, & Burke DM (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. Psychology and Aging, 17(4), 662–676. 10.1037/0882-7974.17.4.662 [PubMed: 12507362]

van Berkum JJA, Brown CM, & Hagoort P (1999). Early Referential Context Effects in Sentence Processing: Evidence from Event-Related Brain Potentials. Journal of Memory and Language, 41(2), 147–182. 10.1006/jmla.1999.2641

Van Berkum JJA, Koornneef AW, Otten M, & Nieuwland MS (2007). Establishing reference in language comprehension: An electrophysiological perspective. Brain Research, 1146, 158–171. 10.1016/j.brainres.2006.06.091 [PubMed: 16916496]

Van Dyke JA, & Lewis RL (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. Journal of Memory and Language, 49(3), 285–316. 10.1016/S0749-596X(03)00081-0

Van Dyke JA, & McElree B (2006). Retrieval interference in sentence comprehension. Journal of Memory and Language, 55(2), 157–166. 10.1016/j.jml.2006.03.007 [PubMed: 18209744]

Verhaeghen P (2003). Aging and vocabulary score: A meta-analysis. Psychology and Aging, 18(2), 332–339. 10.1037/0882-7974.18.2.332 [PubMed: 12825780]

Voss JL, & Federmeier KD (2011). FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. Psychophysiology, 48(4), 532–546. 10.1111/j.1469-8986.2010.01085.x [PubMed: 20701709]

Voss JL, & Paller KA (2006). Fluent Conceptual Processing and Explicit Memory for Faces Are Electrophysiologically Distinct. Journal of Neuroscience, 26(3), 926–933. 10.1523/JNEUROSCI.3931-05.2006 [PubMed: 16421312]

Voss JL, Schendan HE, & Paller KA (2010). Finding meaning in novel geometric shapes influences electrophysiological correlates of repetition and dissociates perceptual and conceptual priming. NeuroImage, 49(3), 2879–2889. 10.1016/j.neuroimage.2009.09.012 [PubMed: 19770053]

Wickelgren WA (1965). Short-Term Memory for Phonemically Similar Lists. The American Journal of Psychology, 78(4), 567–574. JSTOR. 10.2307/1420917 [PubMed: 5839924]

Wilding EL (2000). In what way does the parietal ERP old/new effect index recollection? International Journal of Psychophysiology, 35(1), 81–87. 10.1016/S0167-8760(99)00095-1 [PubMed: 10683669]

Winter B (2020). Statistics for Linguists: An introduction using R. Routledge. DOI: 10.4324/9781315165547

Zhang H, Eppes A, & Diaz MT (2019). Task difficulty modulates age-related differences in the behavioral and neural bases of language production. Neuropsychologia, 124, 254–273. 10.1016/j.neuropsychologia.2018.11.017 [PubMed: 30513288]

Highlights:

- Phonological similarity facilitates encoding for younger and older adults.
- Automatic processes such as spreading activation do not decline with age.
- Phonological similarity causes retrieval difficulty for younger adults.
- Phonological similarity leads to shallow processing in older adults.

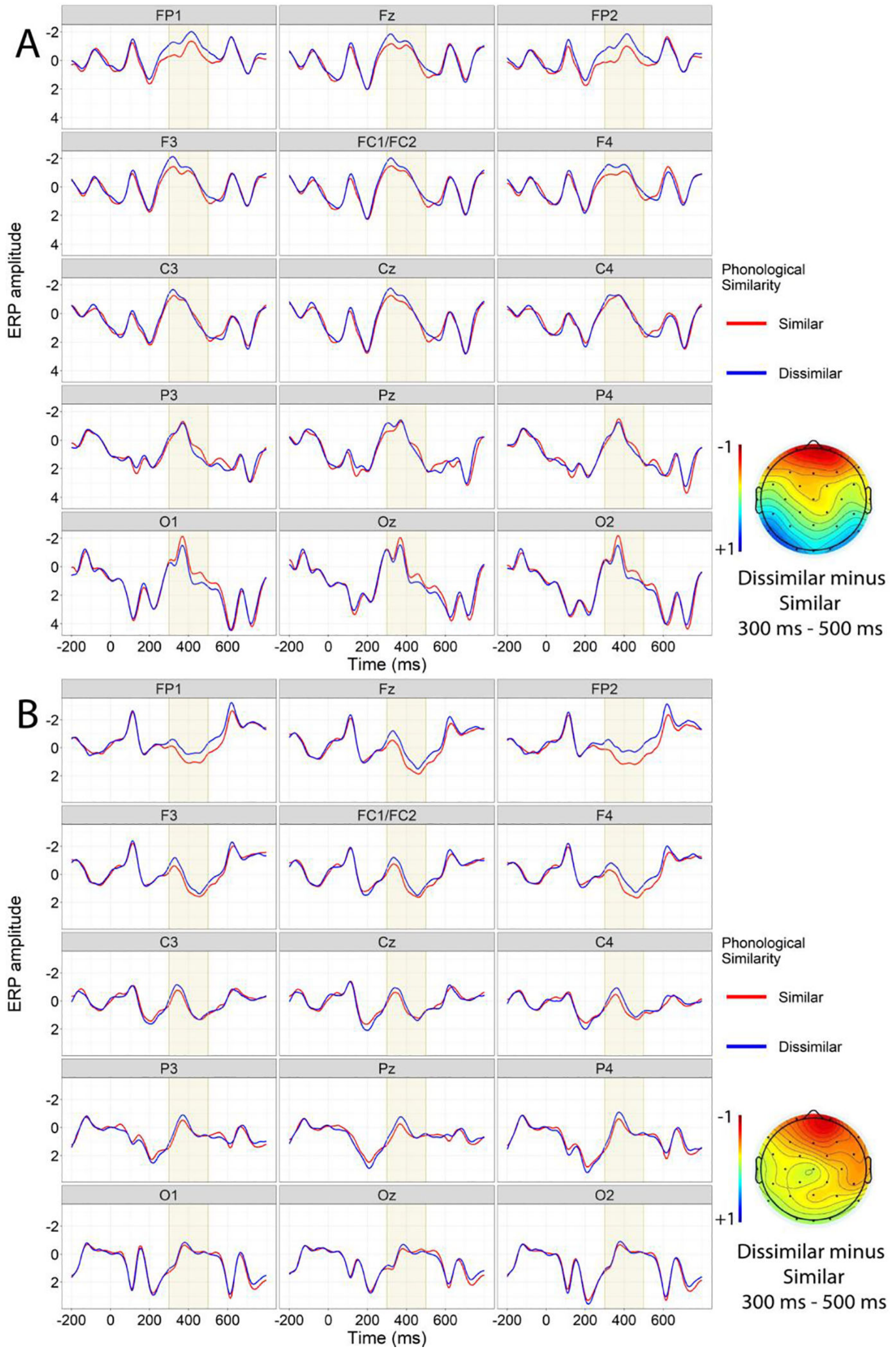

Figure 1.

Grand average brain waveforms of younger (Panel A) and older adults (Panel B) time-locked to NP2 onset. The critical time window is highlighted. Results showed smaller N400 amplitudes (at frontal sites) for NP2s that followed phonologically similar relative to phonologically dissimilar NP1s.

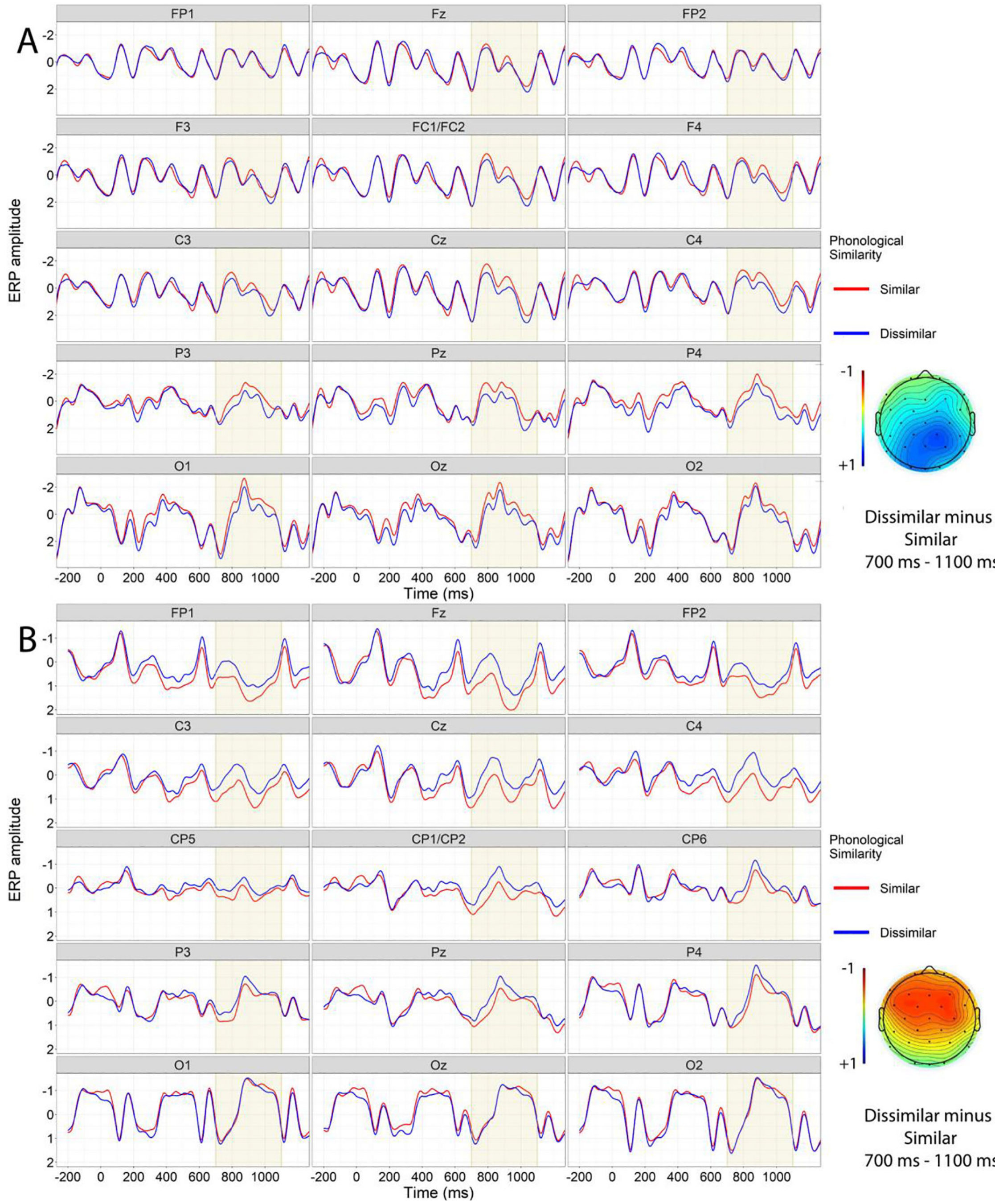

Figure 2.

Grand average brain waveforms of younger (Panel A) and older adults (Panel B) time-locked to pronoun onset. The critical time window is highlighted. Relative to pronouns following phonologically dissimilar NPs, pronouns following phonologically similar NPs caused larger parietal negativities for younger adults, but smaller frontal negativities for older adults.

Table 1.

Results during NP2 encoding in the 300–500 ms time window.

Predictor	Results
Age	$F(1,78) = 11.10$, $p = .001$
Phonological Similarity	$F(1,78) = 3.56$, $p = .06$
Gender Congruence	$F(1,78) = .82$, $p = .37$
Age × Phonological Similarity	$F(1,78) = 1.59$, $p = .28$
Age × Gender Congruence	$F(1,78) = 2.16$, $p = .15$
Phonological Similarity × Gender Congruence	$F(1,78) = .84$, $p = .36$
Age × Phonological Similarity × Gender Congruence	$F(1,78) = 2.14$, $p = .15$
Age × Brain Region	$F(1.20,93.67) = 13.16$, $p < .001$
Age effect at frontal regions	$F(1,78) = 21.9$, $p < .001$
Age effect at central regions	$F(1,78) = 7.39$, $p = .008$
Age effect at parietal regions	$F(1,78) = .04$, $p = .83$
Phonological Similarity × Brain Region	$F(1.32,103.10) = 19.61$, $p < .001$
Phonological Similarity effect at frontal regions	$F(1,79) = 15.46$, $p < .001$
Phonological Similarity effect at central regions	$F(1,79) = 2.37$, $p = .13$
Phonological Similarity effect at parietal regions	$F(1,79) = .70$, $p = .40$

Table 2.

Results during pronoun processing in the 700–1100 ms time window.

Predictor	Results
Age	$F(1,78) = .51$, $p = .48$
Phonological Similarity	$F(1,78) = .05$, $p = .83$
Pronominal Ambiguity	$F(1,78) = .09$, $p = .77$
Age × Phonological Similarity	$F(1,78) = 18.43$, $p < .001$
Similarity effect for younger adults	$F(1,39) = 5.65$, $p =.02$
Similarity effect for older adults	$F(1,39) = 19.10$, $p < .001$
Phonological Similarity × Pronominal Ambiguity	$F(1,78) = .30$, $p = .58$
Age × Phonological Similarity × Pronominal Ambiguity	$F(1,78) = 1$, $p = .32$
Age × Brain Region	$F(1.27,99.42) = 12.52$, $p <.001$
Age effect at frontal regions	$F(1,78) = 6.82$, $p = .01$
Age effect at central regions	$F(1,78) = .14$, $p = .73$
Age effect at parietal regions	$F(1,78) = 2.27$, $p = .14$
Phonological Similarity × Brain Region	$F(1.11,86.21) = 8.72$, $p = .003$
Phonological Similarity effect at frontal regions	$F(1,79) = 3.06$, $p = .08$
Phonological Similarity effect at central regions	$F(1,79) = .36$, $p = .55$
Phonological Similarity effect at parietal regions	$F(1,79) = 4.31$, $p = .04$

Table 3.

Accuracy results for younger adults.

NP Position	Phon. Similarity	Accuracy	Effects
NP1	Dissimilar	88.8%	NP Position = n.s.
NP1	Similar	86.1%	Phon. Similarity = n.s.
NP2	Dissimilar	91.1%	Interaction = n.s.
NP2	Similar	89.0%	

Table 4.

Accuracy results for older adults.

NP Position	Phon. Similarity	Accuracy	Effects
NP1	Dissimilar	89.8%	NP Position = n.s.
NP1	Similar	92.0%	Phon. Similarity = n.s.
NP2	Dissimilar	93.8%	Interaction = significant*
NP2	Similar	91.1%	

Table 5.

Reaction time results for accurate responses for younger adults.

NP Position	Phon. Similarity	RT (ms)	Effects
NP1	Dissimilar	2838	NP Position = n.s.
NP1	Similar	2977	Phon. Similarity = significant*
NP2	Dissimilar	2922	Interaction = n.s.
NP2	Similar	2989	

Table 6.

Reaction time results for accurate responses for older adults.

NP Position	Phon. Similarity	RT (ms)	Effects
NP1	Dissimilar	2931	NP Position = n.s.
NP1	Similar	2915	Phon. Similarity = n.s.
NP2	Dissimilar	2908	Interaction = n.s.
NP2	Similar	2867	

Table 7.

Interpretation bias and associated reaction times for each condition.

Age	Phon. Similarity	NP1	NP2	RT(ms)	Bias results	RT results
Young	Dissimilar	60.1%	39.9%	3807	Age = n.s.	Age = n.s.
Young	Similar	56.9%	44.1%	3906	Phon. Similarity = n.s.	Phon. Similarity = sig*
Old	Dissimilar	60.8%	39.2%	3626	Interaction = n.s.	Interaction = n.s.
Old	Similar	60.2%	39.8%	3800		