

Psycholinguistics and Aging

Michael Ramscar, Eberhard Karls Universität Tübingen - Quantitative Linguistik

<https://doi.org/10.1093/acrefore/9780199384655.013.374>

Published online: 23 March 2022

Summary

Healthy aging is associated with many cognitive, linguistic, and behavioral changes. For example, adults' reaction times slow on many tasks as they grow older, while their memories, appear to fade, especially for apparently basic linguistic information such as other people's names. These changes have traditionally been thought to reflect declines in the processing power of human minds and brains as they age. However, from the perspective of the information-processing paradigm that dominates the study of mind, the question of whether cognitive processing capacities actually decline across the life span can only be scientifically answered in relation to functional models of the information processes that are presumed to be involved in cognition.

Consider, for example, the problem of recalling someone's name. We are usually reminded of the names of friends on a regular basis, and this makes us good at remembering them. However, as we move through life, we inevitably learn more names. Sometimes we hear these new names only once. As we learn each new name, the average exposure we will have had to any individual name we know is likely to decline, while the number of different names we know is likely to increase. This in turn is likely to make the task of recalling a particular name more complex. One consequence of this is as follows: If Mary can only recall names with 95% accuracy at age 60—when she knows 900 names—does she necessarily have a worse memory than she did at age 16, when she could recall any of only 90 names with 98% accuracy? Answering the question of whether Mary's memory for names has actually declined (or improved even) will require some form of quantification of Mary's knowledge of names at any given point in her life and the definition of a quantitative model that predicts expected recall performance for a given amount of name knowledge, as well as an empirical measure of the accuracy of the model across a wide range of circumstances.

Until the early 21st century, the study of cognition and aging was dominated by approaches that failed to meet these requirements. Researchers simply established that Mary's name recall was less accurate at a later age than it was at an earlier one, and took this as evidence that Mary's memory processes had declined in some significant way. However, as computational approaches to studying cognitive—and especially psycholinguistic—processes and processing became more widespread, a number of matters related to the development of processing across the life span began to become apparent: First, the complexity involved in establishing whether or not Mary's name recall did indeed become less accurate with age began to be better understood. Second, when the impact of learning on processing was controlled for, it became apparent that at least some processes showed no signs of decline at all in healthy aging. Third, the degree to which the environment—both in terms of its structure, and its susceptibility to change—further complicates our understanding of life-span

cognitive performance also began to be better comprehended. These new findings not only promise to change our understanding of healthy cognitive aging, but also seem likely to alter our conceptions of cognition and language themselves.

Keywords: Psycholinguistics, learning, aging, language processing, lexical distributions, language change, language use, language production, language comprehension

Subjects: Applied Linguistics, Cognitive Science, Computational Linguistics

Lexical Processing Across the Life Span

The strategy of assessing cognitive or linguistic performance across the adult life span by comparing scores on standardized tests has serious limitations that, unfortunately, have barely begun to be addressed in the literature. These limitations are clearly illustrated by the problem of assessing the lexical processes involved in a simple task like recalling the name of a friend, something that seems to become ever harder for adults to achieve as they age (e.g., Burke et al., 2004; though see James, 2006 for conflicting findings). In studying this question, researchers have typically confronted older and younger adults with tests that involve learning and then recalling names, after which they reported any age-related declines in the performance they observed and theorized as to why the changes they observed had occurred.

However, once we consider the nature and structure of our knowledge of names, and the way that this structure determines its relationship to experience, the inadequacies in this approach soon become clear. Like any other type of word in a language, English given names have a characteristically skewed distribution: In aggregate, English first and last names obey “Zipf’s law” (Zipf, 1949) in that the frequency distribution of names (from most to least popular) roughly follows a power function, which results in a *very* long-tailed distribution (Ramscar et al., 2013c). This means that although a few names are very common as types (Michael, John), most names are not; and the overwhelming majority of names (Winston, Tarquin, Jerome, etc.) are rare. Accordingly, any individual’s knowledge of the total set of names in their community—and indeed, the number of names they know—is highly dependent on how long they have been sampling from the overall name distribution (Ramscar et al., 2014).

The basic statistical properties of names—and of other lexical distributions, because all word types in all the world’s languages appear to obey Zipf’s law—seriously complicates the assessment of name processing across the life span, especially when it comes to determining whether the lexical processes that govern people’s facility with names *decline* as they get older. The long-tailed distribution of given names guarantees that statistically, as we move through life, we learn more and more names. This means that a diligent researcher addressing the question of name memory and aging will be required to control for the fact that the simple task of recalling a name must inevitably get harder across any individual life span (a requirement that, historically, was very difficult to meet). To make matters worse, although people will usually be reminded of the names of their family and friends on a regular

basis (which will help them to remember them), sometimes they will hear a new name only once, or a couple of times at most (in fact, the distribution of names means that, statistically, this is inevitable). It follows that as someone learns a greater number of unique names, the number of different names they know will increase, while the average exposure they will have had to each of the unique names they know is likely to decrease. Meanwhile, although they will meet more and more individuals who share common names, they will also meet more and more individuals with unique names.

These statistical facts make the task of accurately assessing name-recall accuracy far more complex than one might naïvely imagine, simply because they mean that if Anne can only recall names with 95% accuracy at age 60—when she knows 900 first names for 2,000 individuals—it does not follow that she is necessarily worse at processing names than she was at age 16, when she could recall any of only 90 first names for 250 individuals with 98% accuracy. It further follows that testing Anne's ability to recall 20 relatively common names at various points in her life is unlikely to provide any kind of clear, scientific insight into the question of whether her ability to process the many thousands of English names she will be exposed to in her life has declined or even improved. Rather, since the task of assessing her memory for names is complicated by the information-processing differences she faces at any point in time in her life, these differences must—by necessity—be controlled for.

That is, in order to establish whether Anne's memory processes actually decline with age one must: (a), quantify what Mary's knowledge of names is at any given point (because the distribution of names guarantees that the set of names she knows will expand constantly); (b), define a processing model that predicts expected recall performance for a given amount of name knowledge; and (c), independently establish the accuracy of that model across a wide range of circumstances. Unfortunately, despite the fact that researchers have often claimed that declines in memory for names are an inevitable part of healthy aging, not one of the studies of age-related lexical access that gave rise to these claims (e.g., Burke et al., 2004; Evrard, 2002) actually meets these basic requirements.

Moreover, it turns out that the task of quantifying name knowledge is even more invidious than has been described so far. The distribution of English names is not only highly skewed, but has also been changing constantly over the course of the past hundred or so years (i.e., for as long as anyone has been studying our ability to recall names in old age). For a variety of reasons, naming conventions changed markedly in the 20th century. Fewer children were given the most common names, and parents began to choose from an ever-increasing pool of different names. Accordingly, if we assume that recalling a name involves some form of information processing—and to date, accounts of cognitive decline all tend to be couched in processing terms—then researchers interested in name recall are forced to grapple with the fact that U.S. social security records show that the amount of information in the cumulative distribution of U.S. first names has increased *exponentially* over the course of the past hundred years (Ramscar et al., 2014). In concrete terms, this means that the information-processing demands imposed by names have increased so much that a 70-year-old today will be required to process somewhere between 10 and 20 times more information in accessing a name from memory than a septuagenarian of the 1950s would have done. This means that—

for English speakers at least—while our increasing knowledge of the quantitative properties of language enables us to shed quite a lot of light on the complexities behind an apparently simple question like “does our memory for names really decline with age?,” the current reality is that because the factors that influence name memory across the life span present such a rapidly moving target on so many fronts, it is far from clear that we are actually in a position to conclusively answer it.

Similar problems pervade the study of other aspects of lexical processing across the life-span, even apparently simple and straightforward psycholinguistic tasks such as deciding whether a letter sequence is a word or not. As time goes by, people not only encounter more names for other people, but they also encounter more street names, place names, brand names, names for newly invented technology, specialized vocabulary for domains in which they acquire expert knowledge, and so on. Statistically, the distribution of lexical items guarantees that any linguistically active individual will continue to encounter and learn more and more words as time goes by (Keuleers et al., 2015). Unfortunately, however, since a better understanding of these statistical facts has only begun to come about recently, following the development of large electronic corpora and the various tools used in their analysis, it means that historically the increases in lexical knowledge that are inevitable given the long tails of lexical distributions have tended to be missed or massively underestimated by those researching aging. Given this it is perhaps understandable that the answers that researchers have been able to offer to even relatively simple questions, such as what happens to our vocabularies as we age, are often mixed and conflicting: Some studies have claimed that vocabulary knowledge is “preserved,” or even undergoes mild growth as adult age increases (Salthouse, 2014), whereas other studies—and even meta-analyses—suggest that it declines significantly in older adults (Verhaeghen, 2003), with studies even showing significant declines in the vocabularies of men from age 45 on (Singh-Manoux et al., 2012).

For basic statistical reasons, this confused and confusing pattern of results can in fact tell us nothing about the true rate of vocabulary growth across the adult life span (Ramscar et al., 2014). Instead, it serves to further reveal the limitations of repurposing tests developed for measuring the performance of a specific groups in studies of life-span performance: The overwhelming majority of commonly employed vocabulary tests were designed to measure the lexical knowledge of schoolchildren, whose youthfulness ensures that their lexical knowledge is sampled from the top of lexical distributions, and is more homogeneous than that of adults (Keuleers et al., 2015; Meylan & Gahl, 2014). Unfortunately, the long tails of lexical distributions guarantee that vocabulary growth in adulthood is increasingly driven by what individuals learn from their own, idiosyncratic sampling of the tail of the distribution of lexical types, and although this is by far the largest part of this distribution, tests normed on schoolchildren—and then repurposed for studies of aging—are not designed to measure knowledge of it (Ramscar et al., 2013b, 2014).

The force of the statistical facts that determine that linguistically active adults must necessarily come to know more and more words can be appreciated by examining life-span performance on other tasks, such as lexical decision studies, which are more sensitive to knowledge of the whole lexical distribution. The accuracy data for monosyllabic

monomorphemic words in the English Lexicon Project (Balota et al., 2007) is plotted in Figure 1, and it reveals that older adults' accuracy in deciding whether a letter sequence presented to them is a word or not is *far* greater than the accuracy of younger adults, and that it increases across the range of frequencies such that for the lowest frequency words, whereas the average performance of younger adults is at chance, the average accuracy of older adults exceeds 80% (Ramskar et al., 2013b).

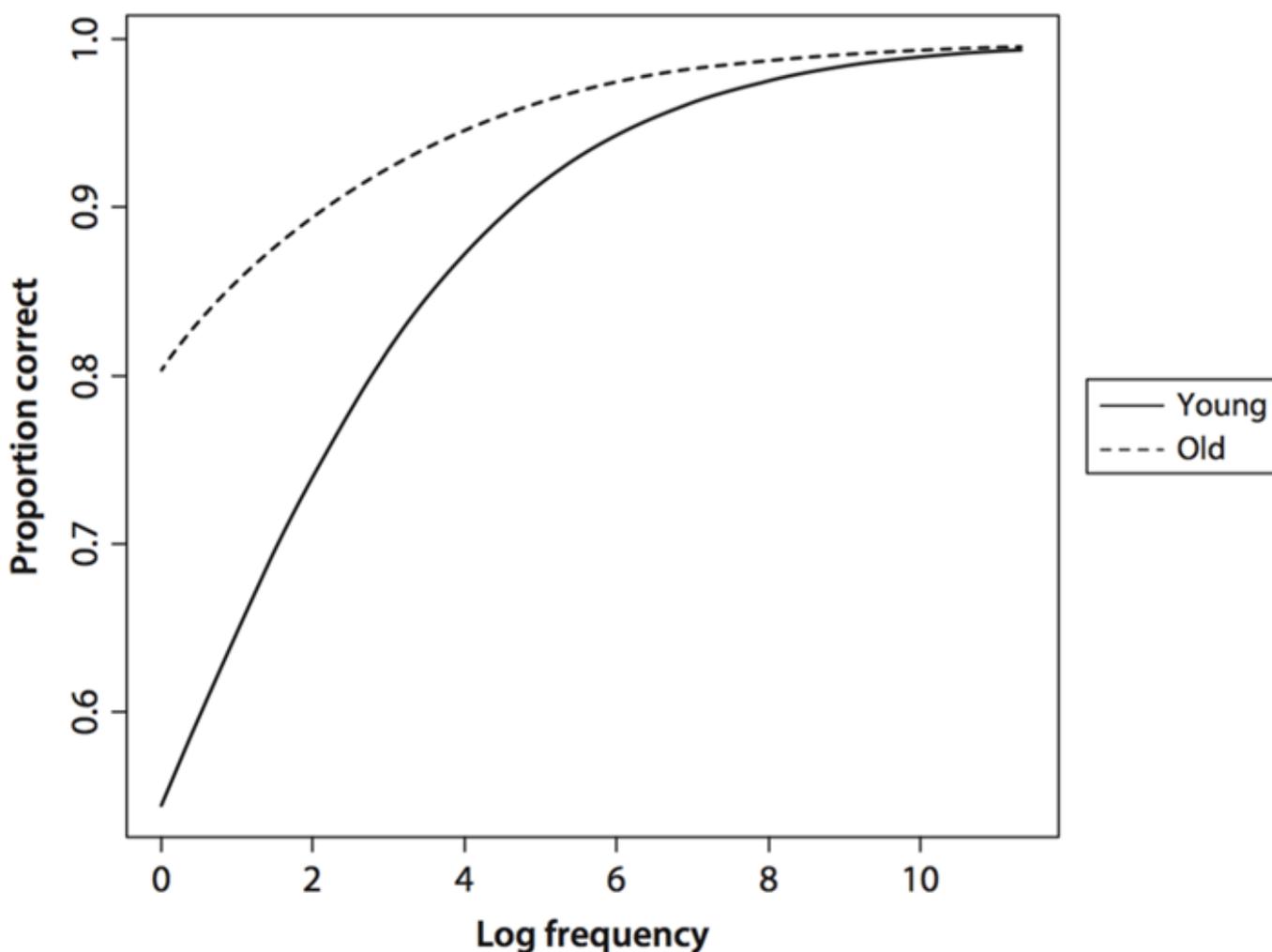


Figure 1. The proportion of correct responses in a visual lexical decision for the young and old subjects in Balota et al. (2007), plotted as a function of log word frequency in a logistic linear mixed model. Old subjects increasingly outperform younger subjects as frequency decreases (Ramskar et al., 2013).

Further confirmation of the nature of vocabulary acquisition and its continued growth across the life span, comes from psycholinguistic measures that *are* specifically designed to measure knowledge in the tail of lexical distributions. Keuleers et al. (2015) employed crowdsourcing to measure lexical decision performance across both a large sample of a population and the full range of its linguistic distribution and found that accuracy continued to increase systematically over the lifetime such that accuracy at age 70 is significantly greater than at

age 60, which is significantly greater than at age 50, and so on (Figure 2). The results of this test are further supported by analyses of spontaneous speech, which reveal older speakers' conversational vocabulary to be richer than that of younger speakers (Gerstenberg, 2015; Meylan & Gahl, 2014).

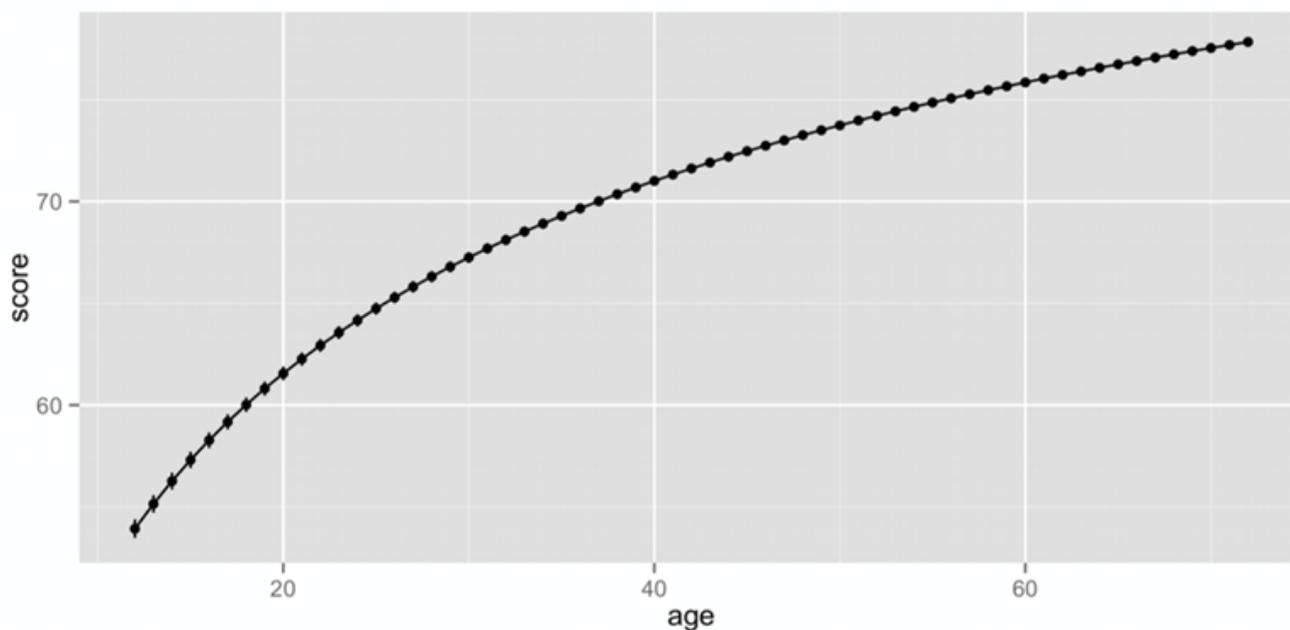


Figure 2. The effect of age on accuracy in a lexical decision task that was designed to measure knowledge of the tail of the lexical distribution (Keuleers et al., 2015). As can be seen, vocabulary knowledge clearly continues to grow across the life span. (Note at least some of the slowing of the growth rate in this plot reflects the fact that this test set is being acquired to near ceiling performance, as opposed to an asymptote in the growth of vocabulary knowledge.)

The results of these studies serve to underline the many challenges that the statistical properties of lexical distributions pose for the accurate measurement of individual vocabulary sizes (such that naïve measures may fail to detect vocabulary growth, or even register spurious declines in vocabulary knowledge), not the least of which is the statistical matter of fact that healthy individuals' knowledge of their language's lexicon will inevitably increase as the time they spend engaging with their linguistic community increases.

This knowledge growth is nontrivial, and it has important—even critical—implications for any serious scientific attempt to understand and model ordinary, healthy cognitive functioning across the life span. Historically, all age-related changes in reaction time and performance on memory tests have been assumed to provide evidence of declines in processing power or capacity. Yet in the light of the fact that adults continue to learn across the life span—as evidenced by their continually increasing vocabulary knowledge—it is clear that these assumptions contradict themselves: If we accept that our minds and brains can be best

understood by treating them as natural information-processing devices—as descriptions of the effects of aging inevitably do (see e.g., Deary et al., 2009; Salthouse, 1988)—then it turns out that a simple fact about information-processing systems is that speed and accuracy are, by definition, a function of the amount of information that is being processed and the capacity of the processing device (Shannon, 1948).

Formally, information-processing systems are not only “digital” in the mundane sense that they make use of binary codes of ones and zeros, but also in the more interesting sense that “information” is defined in terms of the uncertainty that is present in the distribution of a set of discriminably discrete states. The information communicated by a code referencing a given state is a function of its uncertainty within a predefined system, and this means in turn that the amount of information in a system is a function of the number of discrete states discriminated within it, along with the way that these states are distributed. Critically, this allows us to quantify information, along with gains in information (Shannon, 1948): Consider a simple situation in which someone knows four words, enabling them to discriminate between and communicate about four experiences. In this case, a two-bit code (10, 01, 00, 11) would be sufficient for their purposes; however, if their vocabulary size increases, they will now need a larger code.

What this means, put simply, is that if learning and experience serve to increase the number of discriminable states in an adult’s memory—that is, if an adult learns more words—then even allowing for the fact that practice may lead to gains in tests of performance for individual memory items treated in isolation, it follows that the amount of information that must be processed across the entire lexical system must ultimately increase (Ramscar et al., 2014). In children and adolescents, whose brains are still developing, vocabulary increases tend to predict improved processing speeds (Fernald et al., 2013), which suggests that some of the increase in the information load that is accrued from learning may be offset by changes in neural morphology that effectively increase processing capacity (Thompson-Schill et al., 2009), or by changes in the way that lexical knowledge is represented. However, by contrast, maturation seems to bring more stability to lexical representations (Ramscar et al., 2017) and the morphological structure of the brain (Burke & Barnes, 2006), suggesting that linguistic processing capacity may be less subject to change in adulthood.

If the capacity of a processor is constant, this means that when the amount of information that has to be processed is increased, then something—for example, speed or accuracy—will simply have to give. Accordingly, given that it is clear that lexical knowledge must and does increase across the life span, it follows that unless the information gains that result from this are controlled for, it is impossible to determine whether the performance changes on simple lexical processing tasks that are observed as adults age are evidence of declining cognitive capacities, or whether they simply reflect the predictable consequences of learning on information processing.

Studies involving bilinguals further support this contention: When the performance of bilinguals and monolinguals is compared on lexical processing tasks, it seems clear that the increased costs of having to search two vocabularies also increase response times, such that

the response speeds of younger bilinguals in their first language are comparable to those of older monolinguals (Gollan et al., 2008). Moreover, the performance differences observed between bilinguals and monolinguals make clear that slower reaction times in lexical tasks are not, in themselves, evidence of processing *deficiencies*. Rather, it seems clear that becoming bilingual changes the *nature* of the lexical decision task in a way that makes it harder, because it increases the information-processing costs imposed by the task (Johns et al., 2016).

The differences in the speed and accuracy of lexical decisions in younger and older adults described above indicate that the same situation occurs when older monolinguals are compared to their younger counterparts. When younger adults decide that a low-frequency word they don't know is a "nonword," they make what is to them an accurate and correct response. A straightforward consequence of this is that in a lexical decision task, a young adult is confronted with materials that contain—for them—a higher proportion of nonwords than actual words. This will make real words more noticeable to younger adults, providing yet another reason why it is a mistake to assume that response speeds straightforwardly reflect processing speed, or that slower response speeds are evidence of decline.

What these examples hopefully serve to make clear is exactly why it makes little sense to compare the "processing speed" (or accuracy, etc.) of two people on any given task unless one has some idea of what they are processing, how that processing occurs, and what the processing load for each individual involved is. Yet, historically, simple comparisons of "processing speeds" (or accuracy, etc.) represent the norm in studies of life-span cognitive performance (see e.g., Salthouse, 1996). Unfortunately, a simple comparison of the cognitive performance of older and younger adults in the absence of any kind of a processing model is analogous to a comparison of people's strength that considers only the size of the objects they can carry, but not their weight. Accordingly, it seems inevitable that these practices will have led to the drawing of numerous spurious conclusions about the effects of aging on lexical—as well as other—processes.

The Effects of Experience on Putting Words Together

It seems clear that language learning is a lifelong process. Accordingly, determining the effects of age and experience on cognitive processing also requires that the nature of learning itself be taken into consideration, along with the influence it might have on language processing across the life span. It is thus perhaps unfortunate in this regard that for most of the past 50 or so years, many linguists have been convinced that the study of learning has little to offer the study of language (Chomsky, 1959), and as a corollary, for much of this period learning theory had little influence on the cognitive sciences more generally (Ramscar et al., 2013a; Rescorla, 1988). Thus, although simple lexical associative learning tasks have been employed extremely widely across studies of aging, and although learning stands almost alone in being a scientifically well-understood psychological process (integrating a range of findings from animal, behavioral, neuroscience, and computational approaches), it is perhaps

not altogether surprising that the insights of the models that have emerged from studies of associative learning have so far had little influence on the way aging researchers interpret the results of these tests.

Before explaining how important this is to understanding the effects of age on language and cognition, it is worth noting that the most important finding that has emerged from a century of research into associative learning is rather counterintuitive: A vast array of findings support the idea that the best way of characterizing the way we come to “associate” two items in memory is the result of a process that actually serves to increase the degree of discrimination between all the things – such as words – that we already know (Ramscar et al., 2010). This process is best illustrated in relation to a measure of associative learning that, as noted above, has been ubiquitously employed in studies of cognition across the life span, the paired-associate learning task (PAL).

PAL tests are considered to be an ecologically valid way of measuring an individual’s capacity to deal with the demands to learn and recall new information that typify everyday life: People are asked to learn word pairs and then recall one word (the target) when given another (the cue). Some pairs are easy, for example, baby–cries, while others are harder, obey–eagle. As Figure 3 shows, performance on this task gets worse as people get older, a finding that is usually taken to reveal how cognitive function declines across the life span.

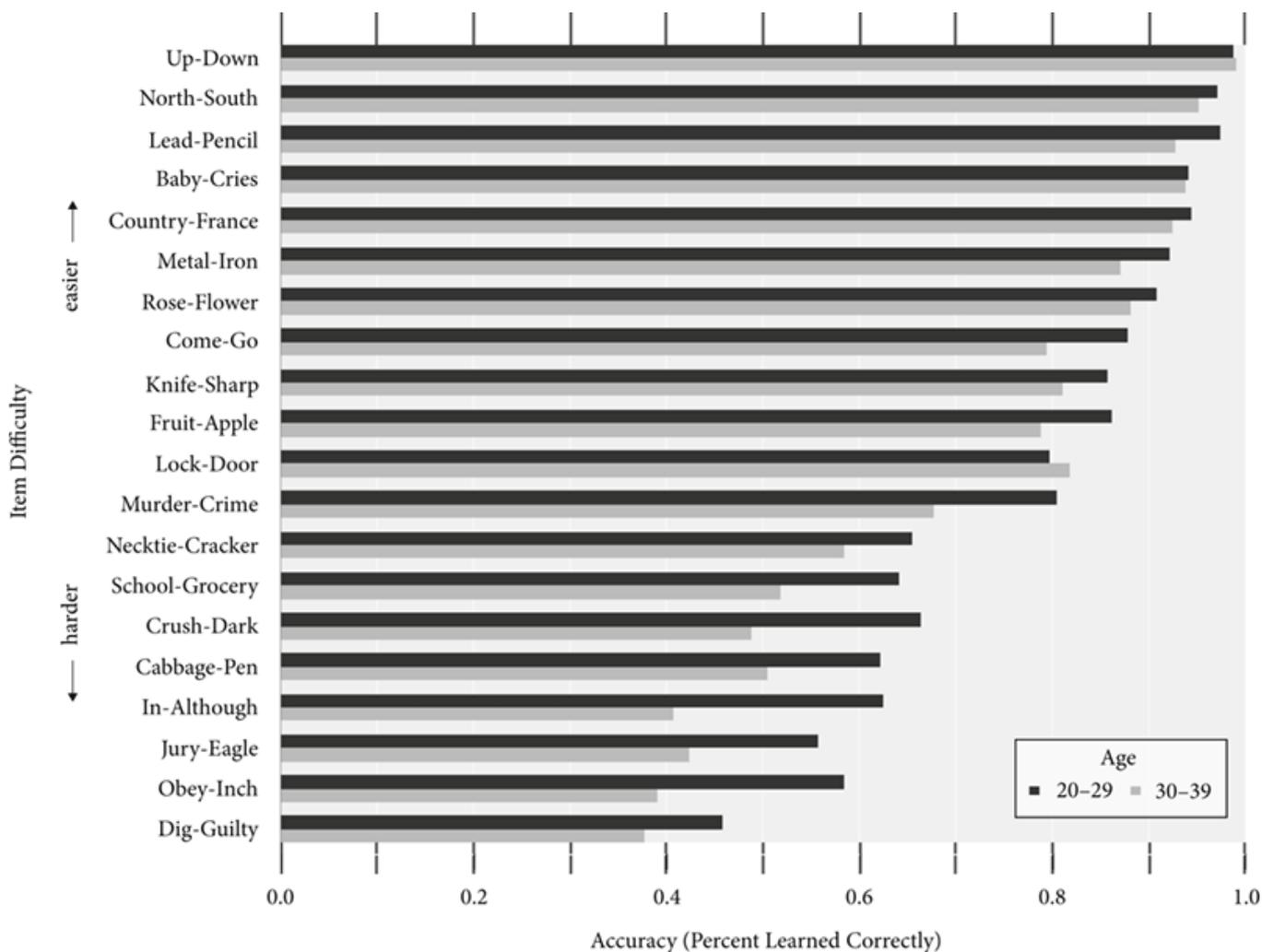


Figure 3. Average by-item performance for 400 adults aged 20-29 and 30-39 (50% females per group) on forms 1 and 2 of the WMS-PAL subtest (desRosiers & Ivison, 1988). The order of items on the y-axis is based on the mean item score across both age groups, with harder items near the bottom of the plot.

Further, given that declines in PAL performance have also been shown to be particularly sensitive to the effect of age on cognition—serving as an excellent predictor of age-related performance decline on other tasks (Rabbitt & Lowe, 2000)—there are some important things to notice about the actual patterns of age-related change that have been observed on this test. First, the biggest change in performance does not come in later adulthood (as naïve theories of cognitive decline might suppose), but surprisingly early, among adults in their twenties and thirties (desRosiers & Ivison, 1988). Second, the pattern of change as age and experience grow is distinctly nonlinear across items, such that hard items become much harder to learn than easier items. Strikingly, while this nonlinear pattern of change in paired-associate learning—which has the beneficial advantage of increasingly discriminating items that usually do go together in English from those that usually do not—is not predicted by accounts of “processing decline,” but is predicted straightforwardly by models of associative learning

itself. Moreover, these models suggest that these changes *ought* to be observed in any situation where the learning process and the English language are held constant, and experience increases (Qiu & Johns, 2020; Ramscar et al., 2017).

To understand why, we need to delve a little deeper into what has been learned about learning since the Russian physiologist Ivan Pavlov conducted his famous experiments on dogs (Pavlov, 1927). Pavlov's results initially led to a theory of learning called associationism: If a cue is present, and an outcome follows, then—as is still widely believed—an animal will inevitably learn to associate the two. While humans, of course, also learn this way, the word “associate” is misleading in this context. The actual learning process is rather more complex than this: It turns out that even explaining the acquisition of “simple association” requires that we stipulate a predictive process that implicitly tests and refines the implicit expectations that prior experience causes it to build. These implicit predictions serve to discriminate those features of the environment that are reliable cues to the events that can be expected to unfold from those features that are less informative, or misleading.

Three main factors have been shown to determine the degree to which associations—and importantly, *dissociations*—between cues and outcomes are learned (Rescorla, 1988): Firstly, the frequency with which a cue co-occurs with an outcome (the *association rate*), typically promotes learning, as naïve associationist theories assume. Cues that reliably predict outcomes are informative, and because learning rewards informativity, high association rates tend to result in associations that are strongly learned. Secondly, however, because learning is sensitive to the information in associations, when cues frequently occur without an outcome (where cues have high *background rates*), they tend to be uninformative about that outcome, and this inhibits the learning of an association. Finally, and similarly, in situations where an outcome is already predicted, learning about a new association does not add information about the outcome, and this too inhibits (or *blocks*) learning about associations. Because learning processes track these three factors, they ensure that cues that promote successful predictions are reinforced at the expense of unreliable predictors, which are increasingly ignored as a result of learning. Indeed, it turns out that the reason that a dog learns to associate a bell with food is not merely to do with the co-occurrence between bell and food, but also because it has learned that any other cues available to it are uninformative in a food context (Ramscar et al., 2010, 2013a).

The discriminative logic that characterizes formal models of the learning processes involved in “associative learning” leads in turn to a revised view of the way that changes in PAL performance should be interpreted. Empirically, it has been shown that adults of all ages find some pairs, such as *baby-cries*, easier than other pairs, such as *obey-eagle*; yet theories that interpret declining PAL performance are notably silent about why this is so, as indeed they are silent about why it is that while PAL learning gets somewhat harder for all pairs as we grow older, it is clear that the latter, harder word pairs become proportionally far more difficult to learn.

The explanation for why some pairs are easier to learn than others is straightforward. We know through experience that babies cry, and the words 'baby' and 'cry' tend to co-occur in numerous contexts. The informativity of this pairing is what makes it easy to produce the word 'cries' when prompted with the word 'baby' when we are asked to learn the word pair *baby-cries*. By contrast, experiences involving *obeying* hardly ever coincide with, or are followed by, experiences with *eagles*, and the words *obey* and *eagle* are both unlikely to co-occur together in context, whereas they are of course likely to co-occur with other words. Accordingly, as our experience accumulates over a lifetime, we will become increasingly confident situations involving obeying will not lead to eagles, and that *eagle* is unlikely to follow *obey* in an English sentence. This means that an older adult instructed to learn a pair like *obey-eagle* in a test is being asked to remember a pairing that experience has increasingly taught them is nonsensical. Indeed, formal learning models predict that as experience increases, this will cause the learning of the association between *obey* and *eagle* to be increasingly and actively inhibited in the ordinary course of affairs (Ramscar et al., 2017).

In other words, learning not only serves to actively associate sensible word pairs; it also serves to actively *disassociate* nonsensical pairings between words that make no or less sense in relation to the experience one accumulates over the lifetime. This is why, from a learning perspective, we would expect that if learning processes continue to function normally across the life span, we ought to expect the PAL task to become harder with age. Critically, until recently it would have been impossible to determine the degree to which the changes seen in PAL performance seen across the life span can be attributed to learning, simply because there was no way of measuring how differences in experience in a phenomenon as complex as a language could influence learning in something as simple as a PAL task. However, the advent of large computational corpora and tools for their analysis now allows us to estimate the expected patterns of connections between words based on the way they occur together in corpora that comprise billions of words of English text and speech.

Using these frequencies to estimate lexical co-occurrences—which can also serve as a proxy for the co-occurrence frequencies of events and experiences—Ramscar et al. (2013b) showed that the patterns of change in PAL performance actually reflect quite closely the distributional factors that ought to determine the increases in the difficulty of hard PAL pairs over a lifetime. These analyses revealed that as healthy adults get older, the degree to which they find individual PAL pairs easy or hard to learn increasingly reflects the actual differences in difficulty predicted by the word co-occurrence patterns of English. Or, in other words, the increased sensitivity to the sensibility of pairings shown by older adults is not evidence of cognitive decline, but rather it simply reflects their ever-increasing understanding of the information structure of the English lexicon (Qiu & Johns, 2020; Ramscar et al., 2013b).

The analyses just described explain why discriminative learning theories (Ramscar et al., 2010, 2013a) predict that PAL performance will decline even when learning capacities are constant; this is simply because cumulative linguistic experience will make meaningless word pairings ever harder to learn. As with the effects of experience on simple lexical processes, the predictions and conclusions about the effects of age and, critically, experience on monolingual PAL can be further tested by examining the performance of bilinguals. Ramscar

et al. (2017) compared the performance of older and younger native German speakers on a German PAL test with that of matched groups of Chinese–German bilinguals, and found that older bilinguals outperformed native speakers in a German PAL test, an advantage that increased with age. That is, when age was controlled for, they found that *less* linguistic experience predicted *higher* PAL scores.

Moreover, this result is actually consistent with many other findings relating to aging and associative learning: For example, Naveh-Benjamin (2000) showed that older adults are worse at learning associations between unrelated “units of information” than when they are meaningfully related, while Castel (2005) showed that older adults are better at associating realistic prices with grocery items than unrealistic prices. Similarly, Old and Naveh-Benjamin (2008) showed that adults encode fewer of the relatively uninformative details provided by background contexts in memory tests as they age. It is notable that although these findings are usually taken to reveal age-related “associative deficits” that are (somehow) lessened when associative information is consistent with the environment, the pattern of learning the informative and neglecting the uninformative seen in these and many other studies is in fact predicted by all functional models of learning (Ramscar et al., 2014; Taler et al., 2019). Moreover, its effects are not only associated with aging: Exactly the same functional pattern of behavior can be seen as infants lose their sensitivity to nonnative phonetic distinctions in the course of learning a native language (Werker & Tees, 1984). What these analyses all have in common is that they describe patterns of data that are entirely consistent with what we know about the processes of associative learning, which in turn suggests that far from declining, adult learning capacities appear to be remarkably consistent across the life span.

Lexical Articulation, Experience, and Language

Although aging research has generally failed to detect the enormous gains in vocabulary knowledge that accompany aging, the reasons for these increases are obvious. Increased age necessarily tends to be accompanied by increased experience, and it follows that the more experiences we learn to discriminate, the more items we have to encode in linguistic signals when communicating. All things being equal, we might also expect that this will cause the code to become more complex in turn. Consistent with this suggestion, Baayen et al. (2017) show that this pattern of change is not only true within an individual’s history, but also seems to be true for the history of individual languages.

Although languages such as Chinese and Vietnamese started out with lexicons dominated by monosyllabic forms (Arcodia, 2007), the phonotactic constraints on what can constitute a usable monosyllabic word soon led to the number of possible forms being exhausted, and the majority of words in Chinese and Vietnamese are now bisyllabic compounds. Compounding is, of course, not the only way of extending a code: The number of discriminable states within a code can also be increased by modulating the fine details of articulation (and hence acoustics) to increase the number of discriminable states in a lexical code. Baayen et al. note that these subphonemic patterns of lexical discrimination are common and are implemented by variations in aspects like tone (familiar from Chinese or Vietnamese), stress (familiar from

English), or acoustic duration, and so on. Meanwhile, research has revealed more subtle subphonemic contrasts in English and in related Germanic languages such as Dutch: Gahl (2008) has shown that the acoustic durations of English homophones such as *time* and *thyme* differ systematically; Kemps et al. (2005a, 2005b) have shown that the acoustic durations of stems in isolation and stems in inflected or derived words also differ systematically; and Plag et al. (2015) describe systematic differences in the acoustic durations of the English suffix *-s* that co-vary with its various morphological functions.

One factor that these findings serve to underline is the complexity of the task of mastering the motor control required in speech articulation. Since practice improves performance for other motor skills, Baayen et al. (2017) employed a variety of techniques to examine whether practice might similarly serve to improve the fine details of articulation across the life span. First, a corpus of electromagnetic articulography (EMA) recordings was analyzed to see whether practice really does influence articulation, and revealed that higher frequency words were indeed articulated more precisely and distinctively than lower frequency words. Then, to examine the effects of subphonemic discrimination over the lifetime, Baayen et al. analyzed the articulation of English vowels spoken by 11 English speakers who had been recorded at regular intervals since childhood, finding that their vowel spaces showed a pattern of expansion with age, such that in each individual, the articulation of vowels became less generic and more lexicalized (Baayen et al., 2017). Or, in other words, Baayen et al. found that with increased age and experience, these adults produced utterances in which each individual word was better discriminated from any other word as compared to the utterances of their younger selves, which suggested that time had improved the accuracy and clarity of their speech.

Putting It All Together: Speech Understanding and Production Across the Age Range

The pattern of change in response to experience described above can also be seen in studies of speech understanding in noisy conditions, or what is commonly known as “cocktail party speech.” Although background noise can often be distracting, it is clear that language users have a remarkable ability to track conversation and scale down unwanted noise. How they do this is not easily explained: Given the basic workings of our ears, it is inevitable that all of the acoustic signals in the environment are initially registered in the brain. Some light on how the brain manages to accomplish this was shed by Mesgarani and Chang (2012), who in a pioneering study directly recorded brain activity from epilepsy patients as they followed instructions to pay attention to one of two competing speakers while ignoring the other. Critically, this revealed that although signals reflecting both speakers did make their way to the auditory cortex, only those signals that appeared to relate to the attended talker seemed to be detected in brain regions usually associated with language processing.

Thus the processes behind “cocktail party” speech understanding follow the same principles as learning, in that they help hearers to tune in to what is informative while ignoring everything else. It is thus important to note that while it was initially assumed—and studies

confirmed—that adults’ ability to understand speech in noisy conditions declines with age, the actual pattern of change is more nuanced: It in fact appears that although adults’ ability to understand the speech of randomly selected speakers in conditions of noise does indeed decline as age increases, there is no discernable deterioration in older hearers’ ability to understand the speech of people they are familiar with as they get older, even when it is presented noisily and out of context (Johnsrude et al., 2013). That is, while older hearers are less able to discriminate words spoken by random speakers, there is no decline in their ability to discriminate words spoken by the speakers they regularly practice hearing.

Age, Experience, and Language

When taken at an abstract level, these results might seem baffling: Why should familiarity with a speaker help a listener decode the contents of a spoken signal comprising a set of phonemic contrasts that were presented under experimental conditions that ensured that they were both highly distorted and unpredictable? The answer, as is underlined by all of the foregoing, is that the aggregated statistical properties of languages, along with the discriminative nature of learning, increasingly tend to render a uniform, abstract idea of “language understanding” untenable when we focus on linguistic knowledge and ability at an individual level, and that this tendency only increases as age and experience grow. Thus, while it is true that people speak “languages,” it is also the case that the distributional properties of any given language—its range of vocabulary, dialects, and registers—also mean that any given individual’s experience of their language is idiosyncratic to themselves and those whom they interact with more frequently, and that these more local aspects of linguistic experience will become more exaggerated as age and experience increase. Or to put it another way, no one person speaks “English” in the abstract; rather, any given individual uses the vocabulary, dialects, and registers relevant to themselves and their local interests to communicate with the people that they interact with, and this means that with experience, people not only become expert in the core of a language—which, statistically, is only a very small part of its total area—but also become expert in talking to the people that they talk to about the things that they talk about using the vocabulary that is best suited for doing so.

Future Directions

Languages change constantly. Historically, *process* was an intransitive verb: Soldiers and monks processed, as they went about their business. Then, in the 20th century, *process* began to be used in transitive ways: meat could be processed in plants, computers began to process data, and scientists began to talk about language processing. This change will have increased the amount of information a typical English speaker would have to process in relation to the verb, because increasing the number of meanings associated with *process* increased the number of constraints and discriminable states associated with it. Thus when researchers now use the modern, transitive use of the word, as they do when talking about the *processing declines* that they associate with aging (Deary et al., 2009; Salthouse, 1996), their employment of this sense of the verb signifies in turn the existence of direct objects—

information, in the form of sentences, words, and so on—that are the subject of this processing, and of course, the existence of something that actually does the processing itself. Yet historically, studies of aging have largely proceeded in ignorance of many of the properties of these objects of processing—in particular, their stimulus and informational properties, along with many of the properties of the learning processes that contribute massively to information processing in the brain. By contrast, when it comes to the study of human communication, the impact of the development of modern electronic corpora and the tools for their analysis have been compared to the invention of the telescope and microscope in the physical sciences (Liberman, in Zimmer, 2012). These developments have already revealed that the statistical properties of language are far richer—and much more counterintuitive—than had been previously imagined. When taken alongside the discriminative nature of learning, these findings also make clear that assessing the true impact of age and experience on language processing poses a far harder problem than researchers have traditionally tended to assume. They have also led researchers in psychology and linguistics to try to seek to understand language processes—and even aging—in quantitative information-processing terms (see Gibson et al., 2019 for a review), as well as to understand linguistic information processing in relation to our ever-growing understanding of the brain’s basic learning mechanisms (Ramscar et al., 2010, 2013a). The initial results from this work indicate that the numerous claims about “processing declines” that have been put forward in the literature are massively overblown (e.g., when learning is controlled for in PAL studies, it appears that processing remains constant as healthy adults age; Ramscar et al., 2013a, 2014, 2017). It is to be hoped that in future research, the requirements for measuring cognitive processing described at the outset of this article will be better satisfied—indeed, given our ever-increasing ability to quantify linguistic experience, it may be that psycholinguistic studies of aging will prove to be particularly well suited for these purposes—so that the question of whether cognitive processes actually do decline in any kind of significant way in the course of healthy aging can be addressed more clearly and more substantively.

Further Reading

Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews Neuroscience*, 7(1), 30–40.

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., Penke, L., Rafnsson, S. B., & Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, 92, 135–152.

Johnsrude, I. S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H. P., & Carlyon, R. P. (2013). Swinging at a cocktail party: Voice familiarity aids speech perception in the presence of a competing voice. *Psychological Science*, 24(10), 1995–2004.

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *The Quarterly Journal of Experimental Psychology*, 68(8), 1665–1692.

- Ramscar, M., Hendrix, P., Love, B., & Baayen, R. H. (2013). Learning is not decline: The mental lexicon as a window into cognition across the lifespan. *The Mental Lexicon*, 8(3), 450–481.
- Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, 6(1), 5–42.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403–428.
- Taler, V., Johns, B. T., & Jones, M. N. (2019). A large-scale semantic analysis of verbal fluency across the aging spectrum: Data from the canadian longitudinal study on aging <<https://doi.org/10.1093/geronb/gbz003>>. *The Journals of Gerontology: Series B*, 75(9), 221–230.
- Wulff, D. U., De Deyne, S., Jones, M. N., Mata, R., & Aging Lexicon Consortium. (2019). New perspectives on the aging lexicon <<https://doi.org/10.1016/j.tics.2019.05.003>>. *Trends in Cognitive Sciences*, 23(8), 686–698.

References

- Arcodia, G. F. (2007). Chinese: A language of compound words? In F. Montermini, G. Boyé, & N. Hathout (Eds.), *Selected proceedings of the 5th Décembrettes: Morphology in Toulouse* (pp. 79–90). Cascadilla Proceedings Project.
- Baayen, R. H., Tomaschek, F., Gahl, S., & Ramscar, M. (2017). The Ecclesiastes principle in language change. In M. Hundt, S. Mollin, & S. Pfenninger (Eds.), *The changing English language: Psycholinguistic perspectives* (pp. 21–48). Cambridge University Press.
- Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, 39(3), 445–459.
- Burke, D. M., Locantore, J. K., Austin, A. A., & Chae, B. (2004). Cherry pit primes Brad Pitt: Homophone priming effects on young and older adults' production of proper names. *Psychological Science*, 15, 164–170.
- Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews Neuroscience*, 7(1), 30–40.
- Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging*, 20, 718–721.
- Chomsky, N. (1959). A review of B. F. Skinner's *Verbal Behavior*. *Language*, 35(1), 26–58.
- Cohen, G., & Faulkner, D. (1986). Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology*, 4, 187–197.

- Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., Penke, L., Rafnsson, S. B., & Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*, 135–152.
- desRosiers, G., & Ivison, D. (1988). Paired-associate learning: Normative data for differences between high and low associate word pairs. *Journal of Clinical Experimental Neuropsychology*, *8*, 637–642.
- Evrard, M. (2002). Ageing and lexical access to common and proper names in picture naming. *Brain and Language*, *81*(1–3), 174–179.
- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*(2), 234–248.
- Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*(3), 474–496.
- Gerstenberg, A. (2015). A sociolinguistic perspective on vocabulary richness in a seven-year comparison of older adults. In A. Gerstenberg & A. Voeste (Eds.), *Language development: The lifespan perspective* (pp. 109–127). Benjamins.
- Gibson, E., Futrell, R., Piandadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, *23*(5), 389–407.
- Gollan, T. H., Montoya, R. I., Cera, C. M., & Sandoval, T. C. (2008). More use almost always means smaller a frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory & Language*, *58*, 787–814.
- James, L. E. (2006). Specific effects of aging on proper name retrieval: Now you see them, now you don't. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *61*(3), 180–183.
- Johns, B. T., Sheppard, C. L., Jones, M. N., & Taler, V. (2016). The role of semantic diversity in word recognition across aging and bilingualism. *Frontiers in Psychology*, *7*, 703.
- Johnsrude, I. S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H. P., & Carlyon, R. P. (2013). Swinging at a cocktail party: Voice familiarity aids speech perception in the presence of a competing voice <<https://doi.org/10.1177%2F0956797613482467>>. *Psychological Science*, *24*(10), 1995–2004.
- Kemps, R., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005a). Prosodic cues for morphological complexity: The case of Dutch noun plurals. *Memory and Cognition*, *33*, 430–446.
- Kemps, R., Wurm, L. H., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005b). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes*, *20*, 43–73.
- Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *The Quarterly Journal of Experimental Psychology*, *68*(8), 1665–1692.

- Mesgarani, N., & Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, *485*(7397), 233-236.
- Meylan, S. C., & Gahl, S. (2014). The divergent lexicon: Lexical overlap decreases with age in a large corpus of conversational speech. In P. Bello, M., Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1006-1011). Cognitive Science Society.
- Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1170.
- Old, S., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*, 104-118.
- Pavlov, I. P. (1927). *Conditional reflexes: An investigation of the physiological activity of the cerebral cortex*. Oxford University Press.
- Plag, I., Homann, J., & Kunter, G. (2015). Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics*, *53*(1), 181-216.
- Qiu, M., & Johns, B. T. (2020). Semantic diversity in paired-associate learning: Further evidence for the information accumulation perspective of cognitive aging. *Psychonomic Bulletin & Review*, *27*, 114-121.
- Rabbitt, P., & Lowe, C. (2000). Patterns of cognitive ageing. *Psychological Research*, *63*, 308-316.
- Ramscar, M., Dye, M., & McCauley, S. M. (2013a). Error and expectation in language learning: The curious absence of mouses in adult speech. *Language*, *89*(4), 760-793.
- Ramscar, M., Hendrix, P., Love, B., & Baayen, R. H. (2013b). Learning is not decline: The mental lexicon as a window into cognition across the lifespan. *The Mental Lexicon*, *8*(3), 450-481.
- Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, *6*(1), 5-42.
- Ramscar, M., Smith, A. H., Dye, M., Futrell, R., Hendrix, P., Baayen, R. H., & Starr, R. (2013c). The "universal" structure of name grammars and the impact of social engineering on the evolution of natural information systems. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3245-3250). Cognitive Science Society.
- Ramscar, M., Sun, C. C., Hendrix, P., & Baayen, H. (2017). The mismeasurement of mind: How neuropsychological testing creates a false picture of cognitive aging. *Psychological Science*, *28*(8), 1171-1179.
- Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*, *34*(6), 909-957.

- Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43(3), 151–160.
- Salthouse, T. A. (1988). The role of processing resources in cognitive aging. In M. L. Howe & C. J. Brainerd (Eds.), *Cognitive development in adulthood* (Springer Series in Cognitive Development, pp. 185–239). Springer.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403–428.
- Salthouse, T. A. (2014). Quantity and structure of word knowledge across adulthood. *Intelligence*, 46, 122–130.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423 & 623–656.
- Singh-Manoux, A., Kivamaki, M., Glymour, M. M., Elbaz, A., Berr, C., Ebmeier, K. P., Ferrie, J. E., & Dugravot, A. (2012). Timing of onset of cognitive decline: Results from Whitehall II prospective cohort study. *British Medical Journal*, 344, d7622.
- Taler, V., Johns, B. T., & Jones, M. N. (2019). A large-scale semantic analysis of verbal fluency across the aging spectrum: Data from the Canadian longitudinal study on aging <<https://doi.org/10.1093/geronb/gbz003>>. *The Journals of Gerontology: Series B*, 75(9), 221–230.
- Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science*, 18(5), 259–263.
- Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging*, 18(2), 332.
- Werker, J., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63.
- Zimmer, B. (2012). New microscopes, new telescopes: A conversation with Mark Liberman about the uncertain future of linguistics. *American Speech*, 87(1), 107–108.
- Zipf, G. K. (1949). *Human behavior and the principle of least-effort*. Addison-Wesley.

Related Articles

Usage-Based Linguistics