Situated Language Processing Across the Lifespan: A Review

Katja Münster[1] & P. Knoeferle[1]

[1] Humboldt-Universität zu Berlin, Germany

Correspondence: Katja Münster, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. E-mail: katja.muenster@hu-berlin.de

Abstract

In this review we focus on the close interplay between visual contextual information and real-time language processing. Crucially, we are showing that not only college-aged adults but also children and older adults can profit from visual contextual information for language comprehension. Yet, given age-related biological and experiential changes, children and older adults might not always be able to link visual and linguistic information in the same way and with the same time course as younger adults in real-time language processing. Psycholinguistic research on visually situated real-time language processing in children and even more so older adults is still scarce compared to research in this domain using college-aged participants. In order to gain more comprehensive insights into the interplay between vision and language during real-time processing, we are arguing for a lifespan approach to situated language processing.

Keywords: situated language processing, children, older adults, lifespan, visual context, age differences, time course, on-line comprehension, eye-tracking

1. Introduction

When communicating with other people, we are often aware of the surrounding visual world. Moreover, when we are talking or listening, we can rapidly make use of this visual contextual information. We can, for example, refer to something we see, or recruit visual information to rapidly facilitate language comprehension and linguistic ambiguity resolution. Indeed, research on adults' language comprehension has demonstrated that non-linguistic visual context contributes rapidly to real-time language processing (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; contra Fodor, 1983). One conclusion from the latter findings has been that the perception of the visual context and of language is often tightly connected and can shape each other's interpretation (see e.g., Knoeferle & Crocker, 2006).

However, when we write about visual context effects, what we really mean is that visual context can influence *younger adults'* language processing, since the majority of our data comes from students between approximately 18 and 30 years of age. In making statements about visual context effects, we thus implicitly generalize from a very small segment of language users (young adults) to the entire population. The fact that we do this is often not made explicit and "young adults" implicitly stands for "all language users". One might argue that such a generalization is largely unproblematic. After all, each young adult was once an infant and will age, suggesting lifespan similarity that might justify generalizing from young adult ages to a model of the "prototypical" language user. However, aging goes hand in hand with biological maturation and decline, as well as with differences in language and world experience and associated mechanisms, all of which could, in principle, modulate (visual context effects in) language processing (cf., Thornton & Light, 2006; and Wingfield & Tun, 2001 for reviews). Thus, demonstrating the existence of rapid visual context effects in young university students arguably amounts to testing the language processing system in specific (perhaps near-perfect) conditions: adults in the prime of their age with excellent cognitive resources, audition, vision and attention, and benefitting from substantial experience in reading, speaking, and in interacting with the world. Likely, testing visual context effects on language processing in these conditions is not equal to testing the same issue in other language users (to the extent that these users differ in their biology and / or experience).

Comprehenders vary, among others, in age and with that in cognitive abilities and / or linguistic and real-world experience. Both age and age-related variation in the comprehender might modulate (or even eliminate?) the rapid integration of visual contextual information during language comprehension. Failing to examine also

less-than-perfect (and even boundary) conditions risks providing impoverished insights into how visual perception (of things and events) contributes to real-time language processing. It might even lead to the formulation of language processing models that are erroneous in that they over-fit specific language users (e.g., young university students) at the expense of others (e.g., children and older adults as well as other segments of the population).

Considering psycholinguistic research on other age groups, we note that research on real-time visual context effects in child language processing is receiving more and more attention (e.g., Borovsky, Elman, & Fernald, 2012; Borovsky, Sweeney, Elman, & Fernald, 2014; Münster, 2016; Trueswell, Sekerina, Hill, & Logrip, 1999; Zhang & Knoeferle, 2012; Zhou, Crain, & Zhan, 2014). What predictions might we make for four-to-five-year-old children? First, they are still in the process of acquiring their native language and they also have less real-world experience than adults. In addition, many of their cognitive functions are not yet fully developed. Children's biological and experiential scaffolding thus likely differs from that of adults. To the extent that this scaffolding modulates visual context effects on language processing, children might not be able to use visual contextual information in the same way (i.e., as quickly and efficiently) as younger adults to facilitate real-time language comprehension.

At the same time, there is good evidence for the importance of the visual environment in child language acquisition. Language acquisition does not happen in isolation and children learn a language by interacting with the (visual) world around them (e.g., Tomasello, 1992, 2000). This acquisition-related use of the visual context implies that children can draw on what they see to enrich and shape the meaning of what they hear. Given the importance of the visual context for language acquisition, we might expect to observe rapid visual context effects also in child language comprehension (see Knoeferle, 2015). Children might, for instance, be able to draw on the visual input in order to overcome difficulties during real-time comprehension such as when they process structurally non-canonical or ambiguous sentences.

By contrast with the blossoming interest in child language processing, the effect of visual contextual information on older adults' (e.g., 60-80 years of age; but also intermediate ages from 30-60) on-line language comprehension remains largely unexplored. Assessing how and with which time course visual contextual information can be used across the lifespan (and thus across age-related biological and experiential states of the language and visual systems) is crucial for developing more accurate models of situated language processing. We can use insights into variation by age (and associated characteristics of the language user) to constrain and fine-tune language processing models. Both age-related biological factors (e.g., cognitive maturation and abilities) and linguistic as well as life experience might modulate visual context effects on language processing, providing insight into what constrains this interaction.

In this review, we will discuss studies investigating the effects of (visual) contextual information on real-time language comprehension in younger adults, children and older adults. On the basis of the reviewed studies we argue that not only younger adults but also children and older adults can use visual contextual information to inform real-time language processing. However, the time course and nature of these effects varies by age (cf., Borovsky et al., 2014; Carminati & Knoeferle, 2013; Münster, Carminati, & Knoeferle, 2014; Münster, 2016; Zhang & Knoeferle, 2012) and might be mediated by different mechanisms regarding the linking between visual and linguistic information (see section 4, e.g., Carminati & Knoeferle, 2013). These results warrant the systematic extension of psycholinguistic research on younger adults to children and older adults.

In order to investigate the nature of the link between visual and linguistic information, we can use the so-called "visual world paradigm" (Figure 1, see Pyykkönen-Klauck & Crocker, 2016). In this paradigm, eye fixations are recorded while participants, for instance, inspect agents depicted as performing different actions towards the middle character (i.e., the patient, Figures 1 & 2) and listen to a sentence. The visual context is thus, in principle, available for the language user while spoken language is processed in real time (e.g., Eberhard et al., 1995; Tanenhaus & Spivey-Knowlton, 1996; Tanenhaus et al., 1995). Because of this, we can investigate the real-time influence of language processing on the visual interrogation of a scene and of visually perceived information on language processing. Drawing conclusions about language processing from eye movements to referents, requires assumptions about the relationship between the eye-tracking data and the cognitive processes they reflect (Knoeferle, 2015). Consider an example presented in Figure 1 (*Den Marienkäfer (object / patient) kitzelt der Kater (subject / agent)*, paraphrased translation: "The ladybug is tickled by the cat."): During "ladybug", people mostly inspect the ladybug. The verb is the first word that can be mapped unambiguously onto the depicted action (and its associated target agent, the cat). Upon hearing *tickles*, participants rapidly fixated the correct target agent (the cat) more than the mouse even before they heard *cat*. Participants visually anticipated (i.e., fixated) the target character before its mention and the depicted actions were the only cue to the role relations

involving the target agent prior to its mention. We can hence assume that the actions facilitated real-time language processing and thematic role assignment (cf., Knoeferle, Crocker, Scheepers, & Pickering, 2005; Knoeferle & Crocker, 2006, 2007; Münster, 2016).

Figure 1. Example of a visual world paradigm setup

Note. A participant inspects a visual display and listens to a scene-related utterance. At the same time, an eye-tracker records her eye movements to the displayed characters (e.g., a ladybug and a cat for a sentence such as *Den Marienkäfer (object / patient) kitzelt der Kater (subject / agent)*, paraphrased translation: "The ladybug is tickled by the cat.").

Most of the findings we review in this paper originate from the visual world paradigm. In section 2 we discuss results on visual context effects in younger adults' language comprehension, followed by corresponding research on child language processing (section 3). In section 4, we review studies on language comprehension in older adults. Based on this review, we conclude that we must pursue a lifespan approach to how visual perception contributes to language processing in order to develop more accurate models of language comprehension.

2. Younger Adults' Use of Visual Contextual Information for Language Processing

The ability to verbally communicate with one another is crucial for social interactions (e.g., Tomasello, 1992). Verbal communication is rapid and incremental (Kutas & Hillyard, 1980, 1984; Tanenhaus et al., 1995; Tyler & Marslen-Wilson, 1977), and most of the time we are not aware of the underlying physical and cognitive processes. We analyze language input as it unfolds and rapidly interpret it compositionally, using lexico-semantic and structural knowledge (Hagoort, 2003; Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003). Thus, knowledge associated with the linguistic input informs the unfolding interpretation of language. In addition to drawing on preceding linguistic contexts for sentence interpretation (Altmann & Steedman, 1988), we can use visual contextual information to ease language comprehension by resolving structural and role-ambiguities in real time (e.g., Chambers et al., 2004; Knoeferle et al., 2005; Tanenhaus et al., 1995).

2.1 Visual Context Effects on Language Processing

One of the earliest visual world eye-tracking studies demonstrated how visual contextual information influences structural ambiguity resolution in so-called garden-path sentences (Tanenhaus et al., 1995). An example of a garden-path sentence is *Put the apple on the towel in the box*. The context shows an apple on a towel, an empty towel, a box, and a pencil. Here, comprehenders can temporarily analyze *on the towel* either as modifying *the apple* (determining its location on a towel) or as part of the verb phrase (specifying its future destination: the empty towel). The destination interpretation is preferred when only one apple is present, resulting in looks to the apple, then to the empty towel (the incorrect destination) and then to the correct destination (the box) during *in the box*.

However, when two apples are present (e.g., one is on a towel and the other on a napkin), comprehenders instructed to *Put the apple on the towel in the box*, must decide which apple to move. This necessity should bias them to interpret *on the towel* as the location instead of the destination of *the apple*. If so, they should look at the apple located on the towel and then directly at the box (and not at the empty towel), which is precisely what happened. From this the authors deduced that younger adults used the visual context to quickly resolve the syntactic ambiguity in favor of the modifier (vs. verb phrase attachment) analysis and correspondingly the location instead of the destination interpretation (Tanenhaus et al., 1995).

In another visual-world eye-tracking study, Knoeferle et al. (2005) demonstrated that participants between 18

and 30 years of age can use the non-linguistic visual context to rapidly resolve structural and thematic role relation ambiguities in German subject-verb-object (SVO) and object-verb-subject (OVS) sentences. The visual scenes depicted two action events (e.g., a princess washing a pirate while a fencer is painting the princess). The spoken sentence played as participants inspected this scene was initially structurally ambiguous and either described the princess-washes-pirate event (in SVO order) or the princess-is-painted-by-fencer event (in OVS word order). Shortly after the verb had mediated one of the two depicted actions, participants either visually anticipated the pirate (if they had heard *washes*) or the fencer (if they had heard *paints*). From the anticipation of a patient (the pirate in SVO sentences) or an agent (the fencer in OVS sentences), the authors deduced that listeners had assigned a thematic role to the initially role-ambiguous noun phrase *the princess*. Participants could thus use the thematic role information provided by the visual context to resolve the sentence initial structural and thematic role-ambiguity even before the utterance permitted structural disambiguation (at the case-marked determiner of the second noun phrase).

Younger adults can thus use visual contextual information to resolve structural ambiguities. In addition, they can exploit information from the visual context to quickly and incrementally resolve semantic ambiguities. Sedivy, Tanenhaus, Chambers, & Carslon (1999), for example, presented students with workspaces of four real world objects (e.g., a yellow bowl, a yellow comb, a pink comb and a metal knife). While participants inspected the workspaces, they listened to instructions that asked them to touch a target object. These instructions involved a referential expression that included an adjectival modifier, i.e., *Touch the yellow comb*. In half of the trials participants would see two objects of the same color (e.g., a yellow bowl and a yellow comb), hence rendering the verbal instructions ambiguous up to the final noun. In the other half of the trials all objects had distinct colors, hence the visual context was unambiguous with regard to the verbal instruction. The eye movement data indicated that participants quickly and incrementally interpreted the spoken instructions with the help of the objects in the workspace. In the unambiguous conditions, they anticipated the target object already when hearing the modifying color adjective. In the ambiguous condition, however, participants looked at the target object only as they encountered the noun referring to the object.

Furthermore, it seems that younger adults' visual attention during real-time sentence comprehension is boosted more by information depicted in a scene than by what they believe is stereotypical. In a visual-world eye-tracking study, Knoeferle & Crocker (2006) presented participants with clip-art scenes depicting action events. Participants listened to OVS sentences containing a verb mediating two potential agents: One agent was a stereotypical agent of the verb (e.g., a detective for *spying*) while the other agent was not stereotypical (e.g., a wizard for *spying*) but was depicted as performing the action mentioned by the verb (*spying*). Thus, the student participants had to choose between two agents each mediated by a distinct thematic role cue. The eye-movement data suggested that participants were able to use each of the two cues when it uniquely mediated a target agent but preferentially relied on the non-stereotypical action depiction when the two cues conflicted. In the conflicting condition, they fixated the non-stereotypical agent depicted as performing the sentential action more shortly after the verb than the stereotypical agent who performed an unmentioned unrelated action. Participants hence preferred to rapidly rely on the depicted event information more than on their stereotypical thematic role knowledge during comprehension (Knoeferle & Crocker, 2006).

Another study looked at the influence of two different kinds of visual contextual information (action depictions and speaker gaze shifts) on sentence processing in younger adults (Kreysa, Knoeferle, & Nunnemann, 2014). In a visual world eye-tracking study, participants watched videos of a speaker uttering sentences about two of three Second Life characters (e.g., *The waiter congratulates the millionaire in the afternoon*). Actions relating to the sentential verb were versus weren't depicted (by means of an object) and the speaker was either visible, inspecting the characters, or obscured. These cues, when available, were presented just after the onset of the verb, and thus participants, in principle, could use them to anticipate the upcoming patient (*the millionaire*) even before its mention. For instance, upon hearing the verb *congratulates*, participants would see balloons appearing between the waiter and the millionaire (representing *congratulates*) permitting them to anticipate the patient of the verb action. Alternatively, the participants would see the speaker shift her gaze to the millionaire just after the onset of *congratulates,* which, together with the verb, could, in principle, also permit them to anticipate the patient. Results indicated that both the action depiction and speaker gaze were effective in that listeners used them to visually anticipate the upcoming subject / patient in the scene before its mention, thus facilitating real-time language processing.

In summary, this (by no means exhaustive) list of findings suggests that younger adults make extensive use of visual context information for anticipating soon-to-be-mentioned objects and characters (an effect which has been characterized as "facilitative" and "preferential", e.g., Knoeferle, 2015). Based on these and related

findings, psycholinguistic research has characterized language comprehension as expectation-driven, interactive, and as temporally closely coordinated with visual perception.

While this may be an accurate characterization of young adults' language processing, we should consider the characteristics of the sample from which the reviewed findings originate: College-aged adult native speakers are on average highly proficient language users, with 20 to 30 years of real-world experience and associated knowledge, and at the height of their cognitive abilities. Considering this background, it is perhaps unsurprising that they can perceive and interpret extra-linguistic visual information rapidly, link it to related linguistic input within a few hundred milliseconds, and use both visual and linguistic information compositionally to generate expectancies about upcoming linguistic input, thus facilitating language understanding. Delays in young adults' visual anticipation seem rare, but have been observed in studies examining pragmatic processes (e.g., the computation of scalar implicature, Huang & Snedeker, 2008), the resolution of local structural and thematic role ambiguity, and prosodic effects on the comprehension of non-canonical sentences (Weber, Grice, & Crocker, 2006; see Knoeferle & Guerra, 2016 for relevant discussion). Perhaps then, such delays are negligible in characterizing language comprehension processes? For college-educated young adults as a sample, the relatively few instances in which delays seem to emerge might be negligible. But in (a lifespan approach to) modeling language comprehension more comprehensively, we should also assess the age-related (perhaps biologically and experientially motivated) variation in facilitative visual context effects. Below we do just that by asking to which extent children and older adults can understand (canonical and non-canonical) linguistic input in visual contexts.

3. Children's Use of Visual Contextual Information for Language Processing

How does young adults' mostly rapid comprehension compare to children's language comprehension and visual anticipation of upcoming objects? At first glance, the rapid processes observed in young adults emerge early but gradually during infancy. For example, 3-year-olds but not 2.5-year-olds can quickly establish reference using informative color adjectives. In a visual-world eye tracking study by Fernald, Thorpe, & Marchman (2010), infants heard sentences such as *Which one is the blue car?*. The visual context presented was either informative regarding the color adjective (i.e., a blue and a red car) or not (i.e., a blue car and a blue house). The older infants fixated a target object (i.e., car) quicker when the adjective was informative compared to when it was not. Hence, 3-year-olds correctly and rapidly interpreted the color adjective just like younger adults would and anticipated the correct target object. Younger infants, however, struggled to incrementally interpret the sentence and waited for the noun to unambiguously establish reference (Fernald et al., 2010).

Nevertheless, by the age of four children have a robust, basic understanding of their native language. When communicating, they seem to comprehend most of the linguistic input rapidly, suggesting they have acquired knowledge of both vocabulary and compositional structure (Snedeker, 2013). Moreover, already by the age of two, children anticipate and look more at a visually depicted object, such as a cake, upon hearing a constraining verb, i.e., *eat* than at inedible distractor objects in the scene (Mani & Huettig, 2012). Furthermore, 3-5-year-olds show a similar eye-movement pattern and time course as adults during the real-time processing of canonical imperfective and perfective sentences. Children fixated pictures depicting an ongoing (vs. completed) event as quickly as adults when hearing an imperfective (vs. perfective) sentence and vice versa. Hence, even 3-year-olds could—just like younger adults—benefit from the visual context for the real-time processing of grammatical aspect in canonical sentences (Zhou et al., 2014).

However, considering that some delays in situated language processing have emerged even in college-aged adults, we might expect to observe similar or even more pronounced variation in children's use of the visual context for language processing (possibly due to differences in their biological or experiential scaffolding). Few studies have so far investigated to what extent children can rapidly draw on visual contextual information to facilitate the processing of challenging non-canonical or garden-path sentences (e.g., Münster, 2016; Trueswell et al., 1999; Zhang & Knoeferle, 2012). They suggest that even though visual context effects may vary with age among others, they can, in some cases, boost children's processing of challenging syntactic structures.

3.1 Processing Canonical Subject-Verb-Object Sentences

For canonical subject-verb-object sentences, children (Mani & Huettig, 2012; Nation, Marshall, & Altmann, 2003), just like adults (e.g., Kamide et al., 2003), can rapidly anticipate the upcoming mention of objects given a supportive linguistic and visual (clipart) context. When 10-11-year old children listened to sentences like *Jane watched her mother eat the cake* and the visual context showed only one edible object among distractor objects, children launched eye-movements to the only edible object in the scene before having heard that object's name (i.e., already during *eat the*, Nation et al., 2003). For the comprehension of canonical sentences, these results suggest similarities in the anticipation of an expected target object by children and college-aged adults. A further

study on 3-10-year-olds' and adults' inspection of (four) pictures during the presentation of an utterance about familiar event relations (*The pirate hides the treasure*) corroborates this insight (Borovsky et al., 2012). One of the four pictures depicted the grammatical object of the sentence (the treasure); another related to the pirate (e.g., a ship), a third to the action (e.g., hiding a bone), and a fourth was unrelated (e.g., a cat). Participants' anticipation speed of the treasure upon hearing *hides* did not differ by age. Recent research has extended these results to the anticipation of targets based on single-shot learning of previously unfamiliar event relations. Borovsky et al. (2014) described novel relations between familiar animals and objects, such as *The lion flies the kite*. Participants first listened to a short story introducing the actors and their actions while inspecting related comic strips. When they next listened to the target sentence (*The lion flies the kite*) and inspected a set of four objects, both 5-10-year-olds and adults (but not 3-4-year-olds) made more anticipatory fixations to the target object (e.g., the kite) compared with distractor objects during *flies*, mimicking the gaze behavior observed with more familiar events.

While the time course of anticipatory gaze did not appear to vary substantially by age, more detailed analyses (Borovsky et al., 2012; Nation et al., 2003) did reveal some variation in children's comprehension and eye-movement behavior as a function of their comprehension skill, production vocabulary, and age. For instance, 2-year old children's anticipation correlated with their production vocabulary such that only skilled producers anticipated the cake (Mani & Huettig, 2012). In addition, less skilled comprehenders at ages 10-11 fixated the target object in *Jane watched her mother eat…* more often and for a shorter duration than more skilled comprehenders (Nation et al., 2003; but note that children's comprehension skill did not modulate how rapidly they anticipated the cake) Likewise when Borovsky et al. (2012) controlled age-specific vocabulary size, participants with a large (vs. small) vocabulary inspected the target earlier. Thus, vocabulary size appeared to be a more reliable predictor of anticipatory looks than participant age, at least for the processing of canonical linguistic structures. For single-shot learning, Borovsky et al. (2014) observed age-related variation in target anticipation. Three-to-four-year-olds (unlike younger adults and 5-year-olds) did not launch anticipatory fixations to the target (vs. distractors) upon hearing *The lion flies* but only fixated the target (i.e., the kite) as it was mentioned. Perhaps the younger children favored lexical-referential instead of compositional processing of the linguistic and visual input, leading them to inspect objects upon their mention (Borovsky et al., 2014, cf. Zhang & Knoeferle, 2012). Alternatively, or in addition, their arguably more limited experience and cognitive resources delayed visual anticipation.

In summary, these studies revealed that children's eye-movements in a visual display can be mediated by canonical SVO sentences with an adult-like time course. Noticeable variation in participants' visual context integration during real-time language processing emerged for vocabulary size (delays, independent of age), comprehension skill (shorter and more fixations for low skilled than high skilled comprehenders), and age (5- but not 3-4-year-olds anticipated the object of newly-learned event relations in real time).

3.2 Processing Garden-Path and Non-Canonical Object-Verb-Subject Sentences

In addition to these variable contextual influences for the processing of canonical sentences, children's reliance on the visual context differs from that of adults' in disambiguating local structural ambiguity. An example of a prepositional phrase (PP) ambiguity from Trueswell et al. (1999, see Tanenhaus et al., 1995) is *Put the frog on the napkin in the box* (*on the napkin* can temporarily either modify *the frog*, indicating its location, or attach onto the verb phrase, indicating the action destination). The latter is an interesting test case: Children (compared with adults) have not yet reached the peak of their biological and experiential scaffolding. As a result, resolving temporary structural ambiguity—with its likely demands on cognitive resources and linguistic experience—may be more taxing for children than young adults, permitting us to test variation of the comprehension system. However, children can rapidly integrate language and the visual context (see section 3.1), and profit from it in language acquisition (e.g., Yu & Smith, 2012). To the extent that these results generalize, children may also benefit from the visual context when resolving local structural ambiguity during comprehension.

Recall that young adults rapidly resolved the PP ambiguity (*Put the apple on the towel in the box*) towards a location interpretation of the ambiguous prepositional phrase (*on the towel*) when two presented apples differed in location (e.g., only one apple was located on a towel, section 2, Tanenhaus et al., 1995). The young adults inspected the apple located on the towel, and then the box and moved the apple there. By contrast, 5-year-olds arguably failed to consider a visual context containing two frogs differing in one attribute: For *Put the frog on the napkin in the box*, they inspected both frogs (of which one was on a napkin) and an available (empty) napkin. This gaze pattern suggests that they interpreted *on the napkin* as a destination (even when the presence of the two frogs should have biased them towards a location interpretation).

One interpretation of the distinct gaze pattern is that the children ignored the contextual bias (Trueswell et al., 1999). Alternatively, young children pursued a different (visual) referential strategy whereas the adults performed compositional interpretation (see also Borovsky et al. 2014). Consider that the children inspected (and moved) one of the two frogs upon hearing *frog*; and then upon hearing *napkin* they inspected that referent (an empty napkin). While it may be tempting to conclude that the children interpreted the empty napkin as a destination, they may have looked there for a different reason (they were looking for the best matching referent of *napkin*, cf., Zhang & Knoeferle, 2012).

Another structure that is difficult to process for children acquiring a case-marking language is the OVS sentence structure. Determining the thematic agent and patient of a sentence based on case-marked determiners of noun phrases poses problems for children when the constituent order is non-canonical and no further context is present. In a case-marking language such as German, the development of thematic role assignment is ongoing until the age of 7 approximately (Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008). However, as argued above, children do not acquire language isolated from their surroundings. To the extent that this insight from language acquisition extends to real-time language processing, children could, in principle, exploit the visual context to disambiguate who-does-what-to-whom in non-canonical utterances. Two studies investigated precisely this issue (Münster, 2016; Zhang & Knoeferle, 2012; see section 4. for more detail). While inspecting a clipart scene, children and adults listened to unambiguous German OVS (Münster, 2016; Zhang & Knoeferle, 2012) and SVO (Zhang & Knoeferle, 2012) sentences describing the scene in a "who does what to whom" fashion. In half of the trials, participants saw the three characters and the scene depicted the action denoted by the sentential verb; in the remaining trials, scenes only depicted the three characters. Participants' eye-movements were measured while they inspected the scene and listened to the sentence. Post-sentence, participants answered a comprehension question about "who is doing what to whom".

The results by Münster (2016) and Zhang & Knoeferle (2012) suggest that children can indeed use the depicted actions to overcome their processing difficulties for German OVS sentences. The verb denoted the action, likely mediating nearby patients and agents, such that both adults and children visually anticipated the depicted object / patient role filler in SVO and the subject / agent role filler for OVS sentences earlier when an action was (vs. was not) depicted in the scene. This gaze pattern suggested that participants had correctly anticipated the thematic roles with the help of the visual context. Moreover, children's response accuracy for OVS sentences increased reliably when events were (vs. weren't) depicted. For SVO sentences, children's response accuracy exceeded that for OVS independent of the action depiction (Zhang & Knoeferle, 2012). Adults, by contrast, scored at ceiling for both sentence types and regardless of action depiction. Hence, visual contextual information such as the depicted events helped the children to overcome their processing difficulties for the non-canonical OVS structure.

Compared with younger adults, children's act-out and eye-movement behavior in response to situated language processing thus appeared more variable. Whether children can integrate visual information into ongoing language comprehension and how rapidly this happens seems to depend both on their age (and perhaps age-related individual differences in scaffolding such as vocabulary growth among others) and the structure of the linguistic input.

4. Older Adults' Use of the Visual Context for Language Processing

By contrast with children, older adults (approximately ages 60-80) have acquired a lifetime of experiences in language and the world and might hence be on par with younger adults in language comprehension. For instance, hearing as little as 20% of a target word in context permitted both older and younger adults to recognize it (Wingfield, Aberdeen, & Stine, 1991). Moreover, older (vs. younger) adults' story interpretations were more elaborate and enriched in meaning as evidenced by their retelling and interpretation of recently read stories (Adams, Smith, Nyquist, & Perlmutter, 1997). But while (linguistic and life) experience increases across the lifespan, processing speed and executive functions among others decline (e.g., Calder et al., 2003; Mill, Allik, Realo, & Valk, 2009). These biological changes might constrain facilitative visual context effects on comprehension, overall and in real time.

Below we discuss (visual context effects in) older adults' language processing and its variation depending on biological as well as experience-based factors. Although many studies have investigated older adults' language processing, most have focused on purely linguistic contexts. Older adults, for example, have larger vocabularies than younger adults (see Verhaegen, 2003 for a meta-analysis) reflecting vocabulary growth across the lifespan. Moreover, semantic processing is often only slightly impaired or even boosted in older age (e.g., Federmeier, Van Petten, Schwartz, & Kutas, 2003; Laver & Burke, 1993). On the other hand, older age often also goes hand

in hand with decline in working memory (e.g., Just & Carpenter, 1992), a decline that can (e.g., Kemtes & Kemper, 1999) but need not compromise language processing, depending on task demands (e.g., Caplan, DeDe, Waters, Michaud, & Tripodis, 2011). In summary, both biological and experience-based factors seem to constrain older adults' language processing.

In line with this assumption is the finding that highly fluent older adults resemble younger adults' in generating expectations. Federmeier et al. (2002) and DeLong, Groppe, Urbach, & Kutas (2012) measured participants' brain waves while they processed linguistic input in real-time in order to investigate linguistic prediction in older and younger adults. Younger and older participants' electrophysiological brain responses to the manipulated linguistic input, i.e., so called event-related brain potentials (ERPs), revealed differences in their on-line sentence processing. This was evident in the "N400", a negative-going ERP wave, larger in mean amplitude between approximately 300-500 ms after the onset of a semantically unexpected compared with an expected stimulus (Kutas & Hillyard, 1980; for a review see Kutas & Federmeier, 2011). Participants read (DeLong et al., 2012), or listened to (Federmeier et al., 2002), sentences containing a more or less expected word (e.g., *The bakery did not accept credit cards so Peter would have to write a check / an apology to the owner*, DeLong et al., 2012). In addition, DeLong and colleagues measured verbal fluency scores (Benton & Hamsher, 1978). In the reading task, younger (but not older) adults' mean N400 amplitudes increased for *an* compared with *a*. The N400 difference to the article suggested that younger adults had already predicted that a noun preceded by *a* (e.g., *check* but not *apology*) was a contextually likely completion. Older adults did not generate such predictions on average, but those with high (vs. low) verbal fluency scores resembled the younger adults in their N400 differences (to *an* vs. *a*) and associated expectations.

Both biological and experience-based factors might thus cause variation in younger and older adults' language use. However, studies investigating (variation in) visual context effects on older adults' language comprehension remain scarce (extant studies examined mostly linguistic context effects, e.g., Caplan et al., 2011; Federmeier et al., 2002; DeLong et al., 2012). In one visual world eye-tracking study, participants (ages 32 to 77) inspected a visual display of 4 objects while they listened to simple spoken Dutch instructions (e.g., *Look at the piano*, Huettig & Janse, 2015). The gender-marked Dutch determiner singled out a piano as the target (three further objects had a different gender, eliminating them as targets). Given the constraining determiner gender, participants could predict that the piano would be mentioned and launch anticipatory fixations towards it before its mention. These anticipatory fixations were modulated by participants' cognitive abilities but not their age. Participants anticipated the target object more the higher their working memory capacity, independent of their chronological age.

However, chronological age and associated emotion-regulation strategies seemed to matter in visual world research investigating how emotional facial effects modulate the real-time processing of emotionally valenced sentences (Carminati & Knoeferle, 2013, see also Münster et al., 2014). Younger and older adults (18-30 years vs. 60-80 years of age) first inspected a speaker's face that either smiled or looked sad. Next they inspected two photographs side-by-side, depicting events of opposite positive and negative valences and listened to a related positively or negatively valenced sentence. Carminati & Knoeferle assessed to which extent the emotional face affected visual attention during on-line language comprehension (e.g., participants should inspect the positive event photograph more when it is referenced following a positive than negative speaker face). A further question was to which extent this speaker face effect would vary by age. Research on the processing, identification and interpretation of emotion in younger and older adults suggests age-dependent preferences for valenced emotional pictures and faces. While older adults have been shown to be biased (e.g., in looking more and longer) towards positive emotional faces and pictures, younger adults seem to be biased towards negatively valenced faces and pictures (see e.g., Socioemotional Selectivity Theory: Carstensen, Fung, & Charles, 2003; Isaacowitz et al., 2007). Crucially, Carminati & Knoeferle (2013) controlled for age-related differences in verbal fluency and working memory among others such that distinct age-related gaze pattern could not be attributed to associated differences in participants' cognitive abilities. Their results revealed some similarities but also clear age-related differences in gaze behavior: Older and younger adults' fixations to the event photographs did not differ in timing but in quality: Whereas younger adults fixated the negatively valenced event more during comprehension of a negative sentence after having inspected a negative than positive facial expression, older adults fixated the positive event more during comprehension of a positive sentence after having inspected a positive than negative facial expression. Conversely, younger adults did not benefit from a positive speaker face and older adults did not benefit from a negative speaker face in their inspection of the event photographs.

Analyses of the data from Carminati & Knoeferle's study (2013) hence demonstrate variation in visual context influences on sentence processing due to what looks like different mechanisms in how the older and younger

adults use visual contextual information during sentence comprehension. Differences in the gaze pattern as a function of participants' measured cognitive abilities were notably absent.

Lifespan differences also emerged in another visual-world eye-tracking study by Münster (2016) who investigated the effects of two different visual contextual cues on real-time thematic role assignment in German OVS sentences across the lifespan. Three participant groups (4-5-year old children, 18-30-year old younger adults and 60-80-year old older adults) inspected positive emotional facial expressions and depicted action events as two types of visual cues. Participants first saw an emotional speaker face followed by a scene depicting two potential agents and a patient (e.g., agents: mouse and cat; patient: ladybug, see Figure 2). The agents in the scene either each performed or did not perform an action towards the patient. The target agent (the cat), always portrayed a happy facial expression, which was congruent in emotional valence with the (positive) sentence. While looking at the scene, participants listened to a scene-related positively valenced OVS sentence. The sentence described the action (depicted or not) that the target agent was performing towards the patient (e.g., literal translation: *The ladybug* (object / patient) *tickles happily the cat* (subject / agent)).
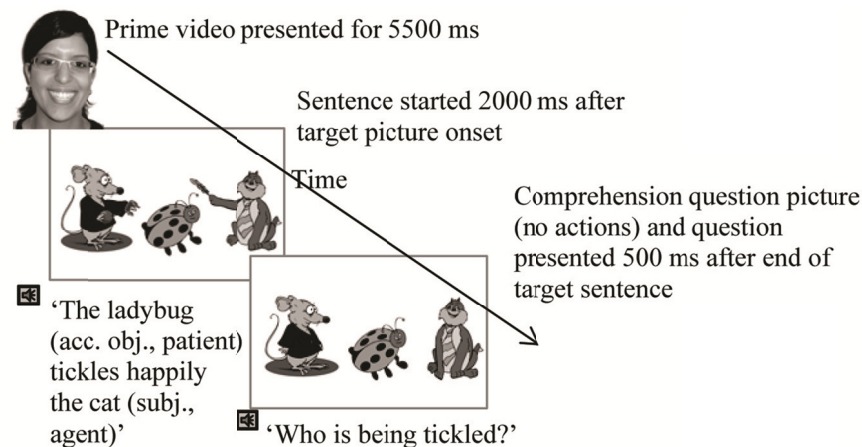

Figure 2. Trial procedure for a trial with a positive emotional speaker prime face and depicted actions (Münster, 2016)

Note. Participants first saw the emotional speaker face in the form of a short video sequence and subsequently encountered the target scene depicting two potential agents (mouse and cat) and a patient (ladybug) either performing (or not) the action mentioned in the sentence. While viewing this scene, participants listened to a positively valenced, scene-related OVS sentence describing the action the target agent (the cat) is performing towards the patient. After the sentence, participants had to answer a comprehension question for which the actions were removed if they had been present (Münster, 2016). Original scenes were in color.

In principle, participants could exploit both of these visual cues (the speaker's facial expression and the action depictions) for anticipating the correct role filler. In the condition in which the positive prime face was congruent in valence with the positive facial expression of the target agent and the positive adverb, participants could link the positive prime face to the positive adverb and the positive target agent's facial expression. In the condition in which the action was depicted, participants could link the verb to the depicted action and hence anticipate its agent. The results crucially showed that all three age groups exploited the depicted action for anticipating the correct agent, facilitating real-time sentence comprehension. However, only the younger and older adults could also recruit the emotional facial expression for thematic role assignment. Visual context effects (of the speaker's face and the actions in older adults; of the actions in children) were delayed in both older adults and children compared with younger adults. Moreover, whereas both adult age groups could use the depicted action in an anticipatory fashion to assign the correct role to the target agent, only younger adults could also use the emotional facial prime predicatively. Older adults only fixated the target agent more given the valence congruent (vs. incongruent) emotional prime once the target agent had been named.

We can hence conclude that older adults show differences in visual attention and language processing compared to younger adults. These differences are likely due to biological factors such as general cognitive decline, slower information processing speed, and perhaps experience-based factors such as differences in (visual) contextual use in older age (see Thornton & Light, 2006; and Wingfield & Tun, 2001 for discussion of age-related changes in information processing). Clearly, we need more research focusing on the use of the visual context in on-line

language comprehension across the lifespan.

5. Conclusions

This review has focused on the facilitative integration of visual contextual information into real-time language processing across the lifespan. We have seen that using visual contextual information for language comprehension is not only beneficial for younger adults (permitting visual anticipation of to-be-mentioned objects), but also for children, especially in more demanding language processing situations. In addition, biological (e.g., executive function) as well as real-world and linguistic experience-based factors (e.g., vocabulary size) likely play an essential role for the use of the visual context during language comprehension. The relevance of these factors becomes particularly evident when looking at language processing in older age. We have reviewed studies suggesting that numerous factors such as verbal fluency and cognitive decline mediate how rapidly and to which extent older adults can process language in real-time. Crucially, this likely also holds for visually situated language comprehension. Just like children and younger adults, older adults showed real-time processing advantages when they could draw on visual contextual information during sentence processing. Just like for children, older adults' reliance on the visual context seemed to be mediated by biological and experience-based factors. However, research on older adults in this domain is still scarce and thus demands greater attention, with the goal of developing accounts of situated language processing across the lifespan.

To conclude, pursuing a lifespan approach to situated language processing has the potential to advance our insights into why and to which extent cognitive abilities and linguistic and real-world experiences mediate the temporally close interplay between visual context information, attention, and language processing. Language processing varies depending on the age and characteristics of the language user. For instance, age-related biological factors and linguistic and life experience can modulate visual context effects on language processing. The effects of these factors can provide insights into what constrains the interaction between visual context information, attention, and language processing, leading to more fine-grained models of situated language processing.

Acknowledgements

References

Adams, C., Smith, M. C., Nyquist, L., & Perlmutter, M. (1997). Adult age-group differences in recall for the literal and interpretive meanings of narrative text. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 52*(4), 187-195. http://dx.doi.org/10.1093/geronb/52B.4.P187

Adams, C., Smith, M. C., Pasupathi, M., & Vitolo, L. (2002). Social context effects on story recall in older and younger women: Does the listener make a difference? *Journal of Gerontology: Psychological Sciences, 57B*, 28-40. http://dx.doi.org/10.1093/geronb/57.1.P28

Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*, 191-238. http://dx.doi.org/10.1016/0010-0277(88)90020-0

Benton, A. L., & Hamsher, K. (1978). *Multilingual Aphasia Examination Manual*. Iowa City, IA: University of Iowa.

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a "theory of mind". *Infant Social Cognition*, 333-373.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417-436. http://dx.doi.org/10.1016/j.jecp.2012.01.005

Borovsky, A., Sweeney, K., Elman, J. L., & Fernald, A. (2014). Real-time interpretation of novel events across childhood. *Journal of Memory and Language, 73*, 1-14. http://dx.doi.org/10.1016/j.jml.2014.02.001

Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia, 41*, 195-202. http://dx.doi.org/10.1016/S0028-3932(02)00149-5

Caplan, D., DeDe, G., Waters, G., Michaud, J., & Tripodis, Y. (2011). Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging, 26*(2), 439.

http://dx.doi.org/10.1037/a0021837

Carminati, M. N., & Knoeferle, P. (2013). Effects of speaker emotional facial expression and listener age on incremental sentence processing. *PLoS ONE, 8*(9), e72559. http://dx.doi.org/10.1371/journal.pone.0072559

Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion, 27*, 103-123. http://dx.doi.org/10.1023/A:1024569803230

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 687-696. http://dx.doi.org/10.1037/0278-7393.30.3.687

DeLong, K. A., Groppe, D. M., Urbach, T. P., & Kutas, M. (2012). Thinking ahead or not? Natural aging and anticipation during reading. *Brain and Language, 121*(3), 226-239. http://dx.doi.org/10.1016/j.bandl.2012.02.006

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008a). German children's comprehension of word order and case marking in causative sentences. *Child Development, 79*(4), 1152-1167. http://dx.doi.org/10.1111/j.1467-8624.2008.01181.x

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research, 24*(6), 409-436. http://dx.doi.org/10.1007/BF02143160

Federmeier, K. D., McLennan, D. B., Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39*(2), 133-146. http://dx.doi.org/10.1111/1469-8986.3920133

Federmeier, K. D., Van Petten, C., Schwartz, T. J., & Kutas, M. (2003). Sounds, words, sentences: age-related changes across levels of language processing. *Psychology and Aging, 18*(4), 858. http://dx.doi.org/10.1037/0882-7974.18.4.858

Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective-noun phrases. *Cognitive Psychology, 60*(3), 190-217. http://dx.doi.org/10.1016/j.cogpsych.2009.12.002

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge: MIT press.

Grossman, M., Cooke, A., DeVita, C., Alsop, D., Detre, J., Chen, W., & Gee, J. (2002). Age-related changes in working memory during sentence comprehension: an fMRI study. *Neuroimage, 15*(2), 302-317. http://dx.doi.org/10.1006/nimg.2001.0971

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience,* 883-899. http://dx.doi.org/10.1162/089892903322370807

Huang, Y. T., & Snedeker, J. (2009). Semantic meaning and pragmatic interpretation in 5-Year-Olds: evidence from real-time spoken language comprehension. *Developmental Psychology, 45*, 1723-1739. http://dx.doi.org/10.1037/a0016704

Huettig, F., & Janse, E. (2015). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience, 31*(1), 80-93. http://dx.doi.org/10.1080/23273798.2015.1047459

Isaacowitz, D. M., Lockenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging, 22*, 147-159. http://dx.doi.org/10.1037/0882-7974.22.1.147

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 98*, 122-149. http://dx.doi.org/10.1037/0033-295X.99.1.122

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133-156. http://dx.doi.org/10.1016/S0749-596X(03)00023-8

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: cross-linguistic evidence from German and English. *Journal of Psycholinguistic*

Research, 32, 37-55. http://dx.doi.org/10.1023/A:1021933015362

Kemtes, K. A., & Kemper, S. (1999). Aging and resolution of quantifier scope effects. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 54*(6), 350-360. http://dx.doi.org/10.1093/geronb/54B.6.P350

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition, 95*, 95-127. http://dx.doi.org/10.1016/j.cognition.2004.03.002

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cognitive Science, 30*(3), 481-529. http://dx.doi.org/10.1207/s15516709cog0000_65

Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language, 57*(4), 519-543.

Knoeferle, P. (2015). Visually situated language comprehension in children and in adults. In R. K. Mishra, N. Srinivasan, & F. Huettig (Eds.), *Attention and Vision in Language Processing* (pp. 57-75). New Dheli: Springer India. http://dx.doi.org/10.1007/978-81-322-2443-3_4

Knoeferle, P., & Guerra, E. (2016). Visually situated language comprehension. *Language and Linguistics Compass, 10*(2), 66-82. http://dx.doi.org/10.1111/lnc3.12177

Kreysa, H., Knoeferle, P., & Nunnemann, E. M. (2014). Effects of speaker gaze versus depicted actions on visual attention during sentence comprehension. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 2513-2518). Austin, TX: Cognitive Science Society.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205. http://dx.doi.org/10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161-163. http://dx.doi.org/10.1038/307161a0

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62*, 621. http://dx.doi.org/10.1146/annurev.psych.093008.131123

Laver, G. D., & Burke, D. M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology and Aging, 8*(1), 34. http://dx.doi.org/10.1037/0882-7974.8.1.34

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—but only for skilled producers. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(4), 843-847. http://dx.doi.org/10.1037/a0029284

Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion, 9*(5), 619-630. http://dx.doi.org/10.1037/a0016562

Münster, K., Carminati, M. N., & Knoeferle, P. (2014). How do static and dynamic emotional faces prime incremental semantic interpretation?: Comparing older and younger adults. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th annual conference of the cognitive science society* (pp. 2675-2680). Austin, TX: Cognitive Science Society.

Münster, K. (2016). *Effects of emotional facial expressions and depicted actions on situated language processing across the lifespan* (Doctoral dissertation). University of Bielefeld, Germany.

Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology, 86*(4), 314-329. http://dx.doi.org/10.1016/j.jecp.2003.09.001

Pyykkönen-Klauck, P., & Crocker, M. W. (2016). Attention and eye movement metrics in visual world eye tracking. In P. Knoeferle, P. Pyykönen-Klauck, & M. Crocker (Eds.), *Visually Situated Language Comprehension* (pp. 76-82). Amsterdam / Philadelphia: John Benjamins Publishing Company. http://dx.doi.org/10.1075/aicr.93.03pyk

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109-147.

http://dx.doi.org/10.1016/S0010-0277(99)00025-6

Snedeker, J. (2013). Children's sentence processing. In R. van Gompel (Ed.), *Sentence processing* (pp. 189-220). New York, New York: Psychology Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632-1634. http://dx.doi.org/10.1126/science.7777863

Tanenhaus, M. K., & Spivey-Knowlton, M. J. (1996). Eye-tracking. *Language and Cognitive Processes, 11*(6), 583-588.http://dx.doi.org/10.1080/016909696386971

Thornton, R., & Light, L. L. (2006). Language comprehension and production in normal aging. *Handbook of the psychology of aging, 6*, 261-287. http://dx.doi.org/10.1016/B9-78-012101-2/64950-015X

Tomasello, M. (1992). The social bases of language acquisition. *Social development, 1*(1), 67-87. http://dx.doi.org/10.1111/j.1467-9507.1992.tb00135.x

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*(3), 209-253. http://dx.doi.org/10.1016/S0010-0277(99)00069-4

Trueswell, J. C., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying online sentence processing in young children. *Cognition, 73*, 89-134. http://dx.doi.org/10.1016/S0010-0277(99)00032-3

Tyler, L. K., & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior, 16*(6), 683-692. http://dx.doi.org/10.1016/S0022-5371(77)80027-3

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging, 18*, 332-339. http://dx.doi.org/10.1037/0882-7974.18.2.332

Wingfield, A., & Tun, P. A. (2001). Spoken language comprehension in older adults: Interactions between sensory and cognitive change in normal aging. *In Seminars in Hearing, 22*(3), 287-302. http://dx.doi.org/10.1055/s-2001-15632

Wingfield, A., Aberdeen, J. S., & Stine, E. A. (1991). Word onset gating and linguistic context in spoken word recognition by young and elderly adults. *Journal of Gerontology, 46*(3), 127-129. http://dx.doi.org/10.1093/geronj/46.3.P127

Weber, A., Grice, M, & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. *Cognition, 99*, B63-B72. http://dx.doi.org/10.1016/j.cognition.2005.07.001

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition, 125*(2), 244-262. http://dx.doi.org/10.1016/j.cognition.2012.06.016

Zhang, L., & Knoeferle, P. (2012). Visual context effects on thematic role assignment in children versus adults: Evidence from eye tracking in German. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual conference of the cognitive science society* (pp. 2593-2598). Austin, TX: Cognitive Science Society.

Zhou, P., Crain, S., & Zhan, L. (2014). Grammatical aspect and event recognition in children's online sentence comprehension. *Cognition, 133*(1), 262-276. http://dx.doi.org/10.1016/j.cognition.2014.06.018