Discourse level processing

Elsi Kaiser

1. Introduction

Since language is fundamentally a communicative tool, a way for people to exchange information and create shared knowledge, an understanding of how speakers and hearers use and comprehend discourse-level information is an important part of a theory of human language processing. This chapter provides an overview of how the visual-world eye-tracking paradigm has been used to investigate the processing and representation of discourse-level information.

In a naturalistic setting, when used for communication, human language rarely consists of individual words or even individual sentences. Rather, it is made up of a sequence of utterances, each with some relation to what came before it and what will come after it. We can refer to this stream of utterances – whether they are produced by one person or multiple people speaking with each other – as discourse. A key aspect of interconnected discourse has to do with *reference tracking*: Comprehenders must be able to figure out which nominal expressions are anaphoric, i.e., refer back to an entity that was already mentioned in the discourse, and which nominal expressions introduce new referents. Furthermore, when faced with ambiguous anaphoric expressions such as pronouns (e.g. 'it', 'he'), comprehenders must have a way of identifying which previously-mentioned referent is the intended antecedent.

Moreover, in addition to knowing when to add a new entity to the discourse model and when to reach back for an already-mentioned entity, comprehenders need to keep track of the relations between entities. For example, if someone claims that Lisa ate a pear, and I know that in reality, she ate an apple, I can correct the speaker by saying 'No, it was an APPLE that Lisa ate" or "No, Lisa ate an APPLE" (capital letters mark prosodic emphasis). Thus, I can use intonational cues or syntactic cues to emphasize that I am correcting something stated in prior discourse, namely the identity of what Lisa ate. To fully understand my intended meaning, the comprehender needs to realize not only that 'apple' is a new referent, but also that the initially mentioned referent 'pear' is being contrasted with 'apple.'

Broadly speaking, a successful model of the discourse-level aspects of language comprehension must include an understanding of the connections between sentences; how people use and comprehend them. In this chapter, we focus especially on the cues that guide the

referential aspects of discourse-level processing. As a starting point for our discussion, we will use the notion of *prominence*. The notion of prominence is often conceptualized in terms of how activated, how accessible a particular referent (or abstract representation) is. It is commonly agreed that referents vary in how prominent/activated they are in the mental model that discourse participants construct, and that the accessibility level of a particular referent changes over the course of a discourse (e.g. Garnham, 2001). A central question is how this abstract notion of prominence is reflected on the linguistic level. Over the course of this chapter, we will see that it is difficult to argue for a fully unified notion of prominence, especially if we consider both syntactic and prosodic representations. To see why, we will examine how visual-world eye-tracking studies have contributed to our understanding of how prominence in the prosodic and syntactic domains guides discourse-level processing – with a particular focus on one of the key aspects of discourse-level processing, namely reference tracking.

The visual-world paradigm is well-suited for investigating questions related to reference resolution, because eye movements to objects in a display are closely time-locked to the potential referents that a listener is considering over time (Cooper 1974, Tanenhaus et al. 1995, see Chapter 3 "Attention in vision and language" by Pyykkönen & Crocker for details). Thus, fixations on different objects or characters provide a measure of what people are considering as potential referents for different linguistic expressions as the sentence unfolds in real time. As a result, the visual-world paradigm provides an excellent means of tapping continuously into on-line referential processing, including temporal changes in terms of which referents are being considered when.

This chapter is structured as follows. In the remainder of this section, I briefly review some theoretical approaches to information structure, and summarize the methodological advantages of using the visual-world paradigm for investigating discourse-level processing (see also Chapter 3 "Attention in vision and language" by Pyykkönen and Crocker for a detailed discussion of the visual-world paradigm).

Then, in Section 2 I discuss the relation between prosodic cues and information structure, and review visual-world studies in this area. In Section 3, I turn to the syntactic level and consider what is known about how information structure is encoded in linguistic structure and how eye-tracking studies have contributed to this area. (For information on visual-world-based work on syntactic and lexical processing more generally, please see Chapter 4 "The role of

syntax in sentence and referential processing" by van Gompel and Järvikivi, as well as Chapter 5 "Reaching sentence and reference meaning" by Engelhardt and Ferreira.) As we will see, the notion of 'prominence' plays an important role both in the prosodic and in the syntactic domain, but prosodic prominence and syntactic prominence have different information-structural correlates: Nouns that are prosodically prominent tend to be associated with newness or contrast, whereas nouns that are syntactically prominent do not show a stable association with newness, and if anything, tend to be more connected to givenness.

In Sections 4 and 5, we will look at the notion of 'prominence' through another aspect of language comprehension that is a central part of discourse-level processing, namely reference resolution. I will review visual-world studies focusing on pronoun interpretation as well as studies on the resolution of other kinds of referring expressions, and conclude that if we conceptualize reference resolution as a process sensitive to how prominently different entities are represented in people's minds, then our view of what factors influence referents' prominence needs to include not only entity-related factors (e.g. givenness) and form-specific information, but also event-related factors such as verb semantics and the coherence relations between different events in the discourse.

Conclusions are presented in Section 6.

It is worth noting that discourse-level processing can be approached from the perspective of the speaker or the listener. In the end, of course, every individual assumes the roles of both speaker and listener. However, since most of the visual-world studies relevant for discourse have approached the question from the perspective of the comprehender,[1] that will also be our perspective here.

1.1 Theoretical approaches to discourse-level information

Compositionally, the meaning of a sentence can be derived from the meanings of the individual words and the way in which they are combined. However, due to the interconnected nature of language, sentences also carry discourse-level information that is separate from the core propositional meaning of the sentence: some parts of an utterance will be new information to the hearer whereas other parts will be old/given information, some parts might contrast with or

---

[1] The visual-world paradigm has also been used to investigate various aspects of language production, e.g. Griffin & Bock, 2000; Gleitman et al., 2007; Hwang & Kaiser, 2014; Van de Velde, Meyer & Konopka, 2014.

contradict prior information, and so on. The terms 'information structure' (Halliday, 1967) and 'information packaging' (Chafe, 1974) refer to the ways in which this information is linguistically realized, for example via prosodic or syntactic means.

There exists a vast literature on the topic of information structure in the domain of theoretical linguistics, with researchers arguing for different kinds of information-structural divisions (e.g. *topic-comment*, Gundel 1974; Reinhart, 1982; *topic-focus*, Sgall & Hajicova, 1977/78; *focus-presupposition*, Chomsky, 1971; Jackendoff, 1972; *theme-rheme*, e.g. Halliday, 1967; *open proposition-focus*, Ward, 1985; see also Vallduví, 1990 for a tripartite division). Generally speaking, the various approaches share the fundamental intuition that some part of every utterance connects to something that the listener already knows, and another part provides new information about this familiar entity or event. In other words, each utterance has a topic (i.e., the thing that the sentence is about) and conveys new information about that topic. The listener is faced with the task of identifying which part of the sentence is the topic, and correctly adding the new information about the topic to her mental discourse model.

What kinds of linguistic cues are available to help comprehenders successfully navigate this process of reference tracking and information-updating? After discussing the methodological appropriateness of the visual-world paradigm in the remainder of this section, in Sections 2 and 3 we explore some of the ways in which the information-structural properties of an utterance can be encoded in the linguistic signal, focusing on prosodic and syntactic cues. As will become clear, the notion of prominence plays an important role in signaling information structure in both the syntactic and the prosodic domains.

1.2 Suitability of the visual-world paradigm for discourse-level research

Psycholinguistic research on discourse-level processing is subject to a number of methodological constraints, which converge to make visual-world eye-tracking a well-suited method for this area (see Chapter 3 for a more detailed introduction to the visual-world paradigm). For example, the auditory nature of visual-world eye-tracking studies is a key advantage, in light of the importance of prosodic information. Furthermore, thanks to the close time-locking between linguistic input and eye-movements, eye-tracking allows us to investigate the moment-by-moment processing of auditory stimuli with great precision: For example, Dahan et al (2002) showed that pitch accents guide the interpretation of temporarily ambiguous nouns, revealing the

interpretative consequences of pitch accents before lexical distinguishing information has been encountered. The rapidity of human eye-movements also makes it possibility to use eye-movements as a tool to gain insights into comprehenders' expectations, and to explore how different linguistic cues – prosodic and otherwise -- shape people's expectations about what will happen next in the discourse. Additionally, in the domain of pronoun interpretation, the visual-world paradigm provides a way of obtaining continuous information about the different referents that participants are considering as the linguistic stimulus unfolds over time.

Another advantage of eye-tracking stems from the fact that, biologically speaking, eye-movements are low-cost, low-threshold responses (e.g. Tanenhaus 2007). This means that eye-movements are a very sensitive measure and can pick up effects that other methods may not be able to capture (e.g. Allopenna. Magnuson & Tanenhaus, 1998). This is beneficial when investigating discourse-level processing, because many aspects of discourse-level information do not impact the core meaning or grammaticality of sentences and thus their effects may be transient or otherwise hard to detect.

In addition to its high sensitivity, visual-world eye-tracking has the advantage of allowing for simple, relatively naturalistic tasks. Some researchers use free-flowing, interactive 'language game' tasks (e.g. Brown-Schmidt, 2005). Other studies use instruction-based paradigms, where participants carry out auditorily-presented instructions (e.g. Tanenhaus et al. 1995, Ito & Speer, 2008; Watson et al., 2008). Both of these kinds of tasks have the benefit of providing explicit tasks and goals that are construed in a consistent way by all participants. However, while instruction-based paradigms work well in many situations, they place various constraints on the kinds of sentences that can be used. Thus, some experiments have used tasks such as picture verification (Arnold et al., 2000; Kaiser & Trueswell, 2004; 2008), where participants are asked to assess, on every trial, whether the sentence they hear matches the picture. Some researchers use 'passive listening' tasks, where on most trials participants are instructed simply to listen while viewing a scene, but on some trials participants are asked to provide story continuations (e.g. Järvikivi et al., 2005; Pyykkönen & Järvikivi, 2009). Another alternative is to instruct participants to click with the mouse on (the image of) the last-mentioned person or thing in each sentence, which on target trials can be the referent of a pronoun or other referring expression (e.g. Kaiser, Runner, Sussman & Tanenhaus 2009, Kaiser 2009). This approach has the benefits of task-driven eye-movements without being restricted to imperative sentences, and provides two

kinds of data relevant to referential processing: Participants' eye-movements as well as their mouse-click choices.

1.3 Addressing challenges of the visual-world paradigm in discourse-level research

However, like all methods, visual-world eye-tracking has its challenges**.** By definition, the visual-world method relies on the presence of a visual display or object array of some sort. Although the visual display allows this type of eye-tracking to tap into referential processing with speed and efficiency, it also brings with it some complications.

Perhaps the best-known concern is the 'closed set' effect, i.e., the possibility that the nature of the linguistic and cognitive processes is distorted by the visual presence of a limited set of potential referents. However, this concern is alleviated by results from word recognition showing that non-displayed competitors also influence language processing (Magnuson et al., 2007) – in other words, the set is actually not closed. Crucially, as noted by Tanenhaus (2007), "the display changes the interpretation, but it does *not* change the underlying process" (p.316). In work on pronoun interpretation, the nature of visual display could presumably influence the prominence of the entities mentioned in the discourse. Thus, the visual display is another factor that guides the reference resolution process, and should be carefully controlled (e.g. by ensuring all candidate referents have comparable visual salience).

A related concern has to do with potential biases arising from the location currently being fixated. For example, if a person happens to be looking at a particular referent, say Lisa, when he hears the pronoun *she*, will he be more likely to interpret *she* as referring to Lisa, rather than other referents in the scene, simply because he was looking at Lisa while hearing the pronoun? To minimize these kinds of effects, experiments often include a look-away clause between the critical sentence and the test sentence with the pronoun, to encourage people to fixate a neutral location away from the potential antecedents (see ex.(7,8,9,12). If a person is fixating the look-away location right before he hears a pronoun, and then moves his eyes to look at one of the characters in the scene upon hearing a pronoun, this provides a clear indication that hearing the pronoun caused the listener to shift his attention to that referent.

The visual display brings up another possible concern regarding the ecological validity of the visual-world paradigm: Much of the language processing that humans engage in on a daily basis occurs in the *absence* of relevant visual information (e.g. chatting with someone about what

they did over the weekend). However, research by Spivey and colleagues suggests that even in the absence of relevant visual input, we still use the visual space as an additional 'memory database' (e.g. Spivey, Richardson & Fitneva, 2004, see also work by Altmann using the 'blank screen' paradigm, e.g. Altmann, 2004). This suggests that language comprehension in the presence of an array of objects vs. without any depicted objects may not be quite as far removed from one another as one might initially think.

2 Intonational marking of information structure

Having reviewed relevant properties of the visual-world paradigm, let us now consider experiments that used this paradigm to explore the processing of discouse-level cues, starting with the domain of prosody. The term 'prosody' refers to the suprasegmental phonological and phonetic properties of sentences, in particular intonational prominence and prosodic phrasing. It is widely accepted that *intonational prominence* – marked by localized excursions in fundamental frequency (f0), called pitch accents[2] – is closely connected to information structure (e.g. Gussenhoven, 1983; Schwarzschild, 1999; Selkirk 1995; Rochemont, 1986).

To see the relation between prosodic prominence and information structure, consider a sentence such as (1b), where the object is prosodically prominent (as shown by the capital letters). Ex.(1b) is an appropriate response to the question "Who did Tom call?", because the prosodically prominent object is the new information that answers the question. In contrast, a response such as (2), where the subject is prosodically prominent, would sound rather odd as a response to the same question. Intuitively, we can conclude that new information tends to be prosodically more prominent than old/given information, which is often de-accented (Selkirk 1995).

(1)   (a) Speaker A: *Who did Tom call?*
        (b) Speaker B: *Tom called MARY.*
(2)   *TOM called Mary.*

---

[2] Changes in pitch are not the only acoustic correlates of accent/emphasis: Accented syllables normally also have greater amplitude and longer duration than unaccented syllables (e.g. Beckman 1996, Ladd 1996).

Although the existence of a connection between prosodic prominence and information structure is clear, what is less well-understood is whether particular types of prosodic prominence map onto particular information-structural categories. In English, conflicting claims have been made regarding the contributions of two kinds of pitch accents, the 'presentational' accent (denoted with H* in the ToBI notation, e.g., Beckman & Ayers, 1994) and the 'contrastive' accent (denoted with L+H*). An H* pitch accent is characterized by an increase in f0, whereas an L+H* accent has an initial lowering of f0 followed by a sharp rise to a high f0 peak. A fundamental question concerns the nature of the relationship between these accents and information structural notions: Are H* and L+H* categorically different, associated with different information structural categories, or is the mapping between information structure and pitch accents more complex?

According to Pierrehumbert & Hirschberg (1990)'s seminal work on the meaning of pitch accents, the two accents do indeed map onto different information-structural categories: they argue that the H* accent marks *new* information that should be added to the listener's mental model of the discourse – for example, the object 'Mary' in ex(1b) when the sentence functions as a response to the question 'Who did Tom call?' In contrast, the L+H* accent signals contrast, i.e., that "the accented item – and not some alternative related item – should be mutually believed" (Pierrehumbert & Hirschberg 1990:296). For example, in example (3), if speaker B's utterance 'Tom called Mary' is correcting speaker B's claim that 'Tom called Anne,' then the object 'Mary' is marked with L+H*, as it is the correct alternative, and contrasts with the incorrect 'Anne'.

(3)   A: *Tom called Anne.*
        B*: No, Tom called MARY.*

However, although researchers agree with the general idea that new information tends to be accented (prosodically prominent) and old/known information tends to be deaccented (not prominent) (e.g. Bolinger, 1961; 1986; Chafe, 1974 and many others), the specific information structural properties of different accent types such as H* and L+H* are still under debate (see e.g. Büring, 1997; Steedman, 2000).

2.1 Visual-world studies on the real-time processing of pitch accents

Visual-world methodology has revealed itself to be well-suited for investigating both (i) the nature of the relationship between prosodic prominence and information structure and (ii) the time-course of prosodic processing.

In recent years, there has been an increasing number of studies using visual-world eye-tracking to investigate different kinds of pitch accents, in particular ways in which the information carried by different pitch accents influences reference resolution. The original work by Pierrehumbert & Hirschberg (1990) was in the phonetic tradition and did not use psycholinguistic experimental methods. It was primarily based on native speaker judgments, naturally-occurring (corpus) examples, and sentences elicited from native speakers. In one of the earliest psycholinguistic studies on this topic, Dahan, Tanenhaus & Chambers (2002) investigated the claim that accented words are interpreted as referring to new information whereas deaccented works are interpreted as referring to already-mentioned information. (The term 'deaccented' is sometimes used instead of 'unaccented' when referring to the absence of an accent in a place where one would otherwise be expected.) Participants listened to instructions (ex.4) to move objects around on the computer screen. Crucially, the display included pictures of cohort pairs such as 'candy' and 'candle' (which overlap in their initial segments) in addition to non-phonologically-related objects such as 'pear' and 'necklace.'

(4)

Context sentence: *Put the candle below the triangle.* **OR**
                      Put the candy below the triangle.

Critical sentence: *Now put the CANDLE above the square.* **OR**
                      Now put the candle ABOVE THE SQUARE

Dahan et al manipulated (i) whether the target noun (e.g. *candle*) was mentioned in the context sentence (i.e. whether it was given or new information when participants heard the critical sentence), and (ii) whether the target noun in the critical sentence was accented (shown in ex.4 with capital letters) or deaccented. Participants' eye-movements during the critical sentence revealed a rapid sensitivity to the presence vs. absence of accenting. When the target word was given information (i.e., the candle was mentioned in both the context sentence and the critical

sentence), people were more likely, early on, to look at the new unmentioned object (cohort competitor, candy) when the target word was accented (CANDLE) than when it was deaccented (candy). Conversely, there were more looks to the given object (candy) when the target word was deaccented (candy) than when it was accented (CANDY). The timing of these patterns (starting 300ms after the onset of the target word[3]) reveals the rapid effect of prosodic prominence: Comprehenders' referential processing is being guided by prosodic prominence while the acoustic input is still ambiguous between 'candy' and 'candle'. This highlights the benefit of using cohort pairs like *candle/candy*: They make it possible to see how prosodic cues influence people's referential processing before the words are disambiguated by segmental differences. In sum, then, prosodic prominence creates an expectation that the upcoming noun is new information.

The targets in Dahan et al's accented condition were marked with a mix of H* and L+H* accents. Subsequent work by Ito & Speer (2008) investigated the contrastive L+H* accent more specifically. They tested L+H* accents on adjectives as well as nouns (e.g. *First hang the green ball. Now, hang the BLUE ball*) in both felicitous and infelicitous contexts, using a task in which participants heard instructions to decorate small Christmas trees and had to choose the right ornament out of a large set. Participants' eye-movement patterns showed that when they heard an L+H* accent on an adjective, they were able to rapidly anticipate that the upcoming noun contrasts in color with a previously-mentioned noun, compared to trials when the adjective was not contrastively accented. Similar anticipatory findings were obtained by Weber, Braun & Crocker (2006) for contrastively-accented adjectives in German. Related work by Sekerina & Trueswell (2012) on 6-year-old Russian children's processing of contrastively-accented adjectives and nouns found that children's processing is less anticipatory but can facilitated by contexts which make the contrast set salient.

Recent work by Watson, Tanenhaus & Gunlogson (2008) compared H* and L+H*directly to see whether they do indeed map on to different information structure categories as claimed by Pierrehumbert & Hirschberg (1990). The results suggests that the distinction between L+H* and H* is actually not very clear-cut and that their information-structural properties can overlap. Similar to Dahan et al (2002), participants in Watson et al.'s studies

---

[3] As mentioned in Chapter 3, it takes some time to program a saccade. Dahan et al. (2002) note that in their set-up, one does not expect to see eye-movements driven by the target word until 200-300ms after word onset.

followed instructions to move objects around on the screen, and the screen included pictures of the target word (e.g. camel), a cohort competitor (e.g. candle), two unrelated objects well as four locations. Watson et al manipulated whether the initial stressed syllable of the target word in sentence (c) had an H* or an L+H* accent. A sample item is shown in (5). The first sentence creates a two-member set (e.g. camel and dog) and the second sentence makes one member of that set salient (dog). Thus, a contrastive accent on the target word in the third sentence would be expected to trigger looks to the other member of the set (camel), whereas a new accent is expected to trigger looks to the unmentioned member of the cohort pair (candle). Because it takes about 200ms to program a saccade in a visual task of this nature, and because the cohort pairs overlap in their early segments, the expectation is that effects of pitch accents should be clearest in the 200-400ms time-window after the onset of the target word (while the segmental cues are still lexically ambiguous between candle and camel).

(5) a. *Click on the camel and the dog.*
b. *Move the dog to the right of the square.*
c. *Now, move the CAMEL/CANDLE below the trials.*

Participants' eye-movements show that L+H* is indeed interpreted contrastively, but that H* is more ambiguous, in that it triggers consideration of both new and contrastive referents. More specifically, a target word with an L+H* accent triggers an increase in fixations to (phonologically-compatible) members of the contrast set (e.g. camel) from the 0-200ms time-window to the 200-400ms time-window (where 0ms is the onset of the target word) – indicating that L+H* is associated with contrast. However, a target word with an H* accent results in fixations to both the contrastive referent (camel) and the new referent (candle) increasing at about the same rate the 200-400ms windows and the 0-200ms window. Watson et al. note that these findings cast doubt on the view that different accents map to mutually exclusive information-structural statuses, and instead point to an overlap in the domain of interpretation (see also Chen, den Os & de Ruiter 2007 for related work on British English, using both natural and synthetic speech in a visual-world paradigm).

As a whole, visual-world studies provide a means of testing how different kinds of pitch accents are interpreted in real-time. The results so far highlight the rapidity with which prosodic

information is used and also reveal the importance of empirically testing theoretical claims about the 'meaning' of different pitch accents.

3 Syntactic marking of information structure

Having seen that prosodic prominence provides signals about information structure and guides real-time reference resolution, we now turn to how syntactic prominence relates to information structure and reference tracking, by discussing structural cues such as (i) syntactic position (e.g. subject vs. object) and (ii) special syntactic constructions (e.g. passives and clefts). Compared to prosodic cues, structural cues are more distinctive: Whereas pitch accents can be regarded as inherently gradient (e.g., a high pitch accent is defined *relative* to other accents), different syntactic constructions or syntactic positions are categorical (e.g. a sentence either is or is not a cleft). Thus, whereas one might ask whether H* and L+H* are really different accents or just different points on a continuum, in the syntactic domain the differences between different structural choices are usually more clear-cut.

Let us start by considering the relation between syntactic position and information structure. Evidence from psycholinguistic experiments and corpus studies suggests that certain positions are associated with certain information statuses. For example, in English, where the canonical word order is subject-verb-object, given information tends to occur early on in the sentence – in particular in subject position – whereas new information occurs later (Clark & Clark, 1977; Clark & Haviland, 1977; Halliday, 1967, Arnold, Wasow, Losongco, & Ginstrom, 2000, Gundel 1988, Birner & Ward 1998, see also Prince 1992 on the association between subjecthood and givenness). In English, due to the relatively rigid word order, it is unclear whether this given-before-new pattern is best characterized in terms of grammatical position (subjects tend to be old/given) or linear order (referents mentioned earlier in the sentence tend to be old/given). However, languages with flexible word order, such as Finnish, suggest that linear order is what matters: In Finnish, a sentence with an old subject and a new object tends to have SVO order, whereas a sentence with a new subject and an old object tends to have OVS order.[4] We will discuss the case of Finnish in more detail below. The connection between the sentence-initial position and given information receives further support from passives. Passivization is

---

[4] Although a given-before-new preference has been observed in various languages (e.g. Féry et al 2010), it is not absolute and does not hold for all languages or all constructions (see Clifton & Frazier 2004).

often regarded as a means to promote a given, topical patient argument into the sentence-initial subject position (e.g., Myhill, 1992).

The observation that old/given information tends to occur in sentence-initial position points to a connection between *syntactic prominence and givenness:* The subject position is traditionally regarded as syntactically more prominent than other argument positions (on the basis of various syntactic criteria), and in flexible word-order languages, the initial position is often analyzed as the most prominent. The link between syntactic prominence and givenness is an interesting counterpart to what we observed earlier, namely the link between *prosodic prominence and newness*. (In fact, in the syntactic domain, the sentence-initial position is often associated not only with givenness but also with *topicality*, i.e., the topic of a particular sentence, the entity that the sentence is about, tends to be realized in the initial position. However, topicality and givenness are not synonymous, as referents can be given information without being topics. The precise definition of 'topic' is also still a matter of debate. Nevertheless, generalizing somewhat, it seems fair to say that there tends to be a correlation between syntactic prominence and givenness.)

However, the correlation between syntactic prominence and givenness is not perfect. If we construe the notion of syntactic prominence broadly, it also includes non-canonical constructions such as *it*-clefts (ex.6). Intuitively, the most syntactically prominent constituent in clefts is the capitalized noun in the clefted position – but this position is normally associated with *contrastive focus,* not with givenness or topicality.[5]

(6)   (a) *It was TOM who called Mary. [subject it-cleft]*
        (b) *It was MARY that Tom called. [object it-cleft]*

For example, ex.(6b) signals that Mary was the one who Tom called, rather than someone else. Similar to ex.(3), with an L+H* accent, a cleft like (6a) is commonly uttered in a context where the non-focused part is known/given information (i.e., Tom called someone), and the clefted constituent Mary provides new information that contrasts with other potential alternative referents (see Halliday, 1967; Rochemont, 1986; Rooth, 1992). Thus, here we have a situation

---

[5] Corpus studies show that sometimes, the clefted constituent is given information and the new information is in the rest of the sentence, but such clefts have a different pitch accent pattern (Prince 1978, Hedberg 1990).

where syntactic prominence is connected to contrastive focus – and in fact normally also occurs with a contrastive pitch accent.[6] Thus, when the notion of syntactic prominence is broadly construed, it does not map straightforwardly onto a singular information-structural category. Nevertheless, it is clear that syntax can provide cues about the information-structural properties of different entities.

3.1 Visual-world studies on the real-time processing of syntactic cues to information structure

Given that syntactic representations carry discourse-level information, we would like to know whether and with what kind of time-course comprehenders use this information during real-time processing. Using the visual-world paradigm, Kaiser & Trueswell (2004) found that in Finnish, comprehenders extract information from non-canonical word orders very rapidly, and use it to construct expectations regarding the discourse-status of upcoming referents. Finnish has canonical SVO order, but all six configurations of subject, verb and object are grammatical in different discourse contexts. A rich case-marking system disambiguates the grammatical roles of the nouns. Kaiser & Trueswell (2004) focused on SVO and OVS, two orders that provide information about the discourse-status of the referents: SVO order can be used when the subject and the verb are both given information or new information, or when the subject is given and the object is new. Noncanonical OVS order, in contrast, is used when the object is given and the subject is new. It is also worth noting that standard Finnish does not have definite or indefinite articles (*the, a*) and thus the word order plays a key role in signaling the distinction between new vs. given information. Thus, an SV sequence can be followed by an object that is given or new information, whereas an OV sequence is followed by a new subject.

In Kaiser & Trueswell's eye-tracking study, participants heard SVO and OVS sentences preceded by short discourse contexts, and viewed visual scenes depicting the entities mentioned in the stories, as their eye-movements were recorded. The eye-movement patterns revealed discourse-level anticipation: Compared to SVO sentences, OVS sentences showed anticipatory

---

[6] This brings us to the question of whether *it*-clefts and SVO sentences with L+H* pitch accents are information-structurally equivalent. This question is not yet fully resolved, but it has been suggested that *it*-clefts mark the clefted element as contrastive and also indicate that it is the only possible alternative (ex.(ii)), while pitch accenting in canonical SVO (ex.(i)) does not signal this kind of exhaustiveness/uniqueness (e.g. Kiss 1998, Delin & Oberlander 1995, but see Green & Jaggar 2003 for a different view).
(i)   A: Tom called Anne.   B: No, It was MARY that Tom called. #In fact, he also called Kate.
(ii)   A: Tom called Anne.   B: No, Tom called MARY. In fact, he also called Kate.

eye movements to a discourse-new referent at the onset of the second noun, even before participants had enough acoustic information to recognize the word. Data about the time-course of people's processing of SVO and OVS sentences – in particular the anticipatory nature of listeners' eye-movements in OVS sentences – could be clearly observed, thanks to the nature of the visual-world paradigm. As a whole, these results show that not only pitch accents but also cues from word order can be used by comprehenders to anticipate upcoming new information.

In related work at the interface of prosodic and syntactic cues, Weber et al. (2006) conducted an eye-tracking study on German investigating whether intonation can trigger anticipatory looks when the case-marking on the nouns and articles is ambiguous. Participants heard temporarily ambiguous S-V-adverb-O and O-V-adverb-S sentences where the morphological marking on the first noun and its article was ambiguous between a subject and an object – thus, the initial noun+verb sequence did not disambiguate between SV vs OV. Crucially, however, the intonational contour for the two word orders was different: In S-V-adverb-O sentences, the nuclear pitch accent was on the verb, whereas in O-V-adverb-S sentences, it was on the object. The participants heard the sentences (e.g. SVO: *The cat$_{ambig}$ chases possibly the bird$_{object}$*, OVS: *The cat$_{ambig}$ chases possibly the dog$_{subject}$*.) while viewing scenes depicting the three entities (e.g., cat, bird, dog). Thus, looks to the bird can be interpreted as looks to a suitable patient, and looks to the dog can be interpreted as looks to a suitable agent. Eye-movement patterns showed that sentences with SVO intonation resulted in anticipatory looks during the verb and the adverb to the not-yet-mentioned suitable patient (bird). In sentences with OVS intonation, eye-movement patterns showed looks to the suitable patient (bird) during the verb, but this preference disappeared during the adverb as looks to the suitable agent (dog) increased. As a whole, these results show that participants were able to use cues from intonation to disambiguate grammatical roles in the absence of morphological cues, revealing a close interplay between processing on the syntactic and prosodic levels. Visual-world eye-tracking methodology is ideal for observing this, because it allows us to track how listeners interpret the grammatical roles of auditorily-presented nouns in real time.

The anticipatory eye-movements observed in the domain of discourse-level information fit within a broader set of eye-tracking findings indicating that comprehenders anticipate different kinds of linguistic information during real-time processing, including semantic and syntactic information (e.g. Altmann & Kamide 1999, Kamide et al 2003, Knoeferle et al. 2005,

see also Arnold & Tanenhaus 2011 on anticipatory inferences based on disfluencies). The picture that emerges is of a system that is very 'forwarding-looking' and uses a wide variety of cues to make predictions about who or what will be mentioned next.

4 Effects of discourse-level prominence on pronoun resolution

So far, we have looked at how visual-world eye-tracking studies contribute to our understanding of how prosodic and syntactic cues guide real-time processing, especially the task of deciding whether to add a new referent to the discourse model or to go back to an old, already-mentioned referent. If we conceptualize the prosodic and syntactic cues in terms of prominence, what emerges is an interesting contrast: On the one hand, nouns that are *prosodically prominent* (bear pitch accents) tend to be associated with newness – either because they are new referents being introduced to the discourse or because they are contrastive (i.e., a new assertion is being made involving the noun). On the other hand, nouns that are *syntactically prominent* do not show a stable association with newness, and if anything, tend to be more connected to givenness. As we saw in Section 3, when we look across languages, the structurally-prominent sentence-initial position is associated with givenness, but on the other hand, nouns that are syntactically prominent due to clefting can convey new information. Though this picture is somewhat simplified, it already suggests that a broad notion of 'prominence' is not sufficiently nuanced.

In this and the following sections, we will take a closer look at the concept of prominence through the lens of reference resolution. During language processing, in addition to having to determine whether a particular noun introduces a new referent or refers back to an already-mentioned entity, comprehenders are also faced with ambiguous pronouns. A form such as 'she', 'it' or 'they' tells the comprehender that the referent has already been mentioned in prior discourse,[7] but on its own does not provide enough information to identify the intended referent. It is widely assumed that prominence plays an important role in guiding the interpretation of these ambiguous forms, such that reduced referring expressions (like pronouns) refer to the most prominent entities, i.e., entities that are prominently represented in the interlocutors' mental models of the discourse (e.g., Ariel, 1990; Givón, 1983; see also Gundel, Hedberg & Zacharski, 1993 for related work).

---

[7] So-called cataphoric pronouns are an exception, because in this case, the pronoun can precede the initial mention of the antecedent (e.g. "When he came home, Peter turned on the TV").

However, whenever we talk about prominence or accessibility, we are faced with the question of what influences how prominent a particular referent is at a particular point in the discourse, as well as how this can be measured. So far, we have encountered at least two kinds of prominence – prosodic prominence and syntactic prominence – but with different information-structural correlates. Thus, we can ask: Given that pronouns tend to be interpreted as referring to prominent entities, how does the presence of entities that are prosodically vs. syntactic prominent guide the interpretation of subsequent pronouns? More generally, what kinds of factors influence how prominent referents are, and thus how likely they are to be interpreted as antecedents of subsequent pronouns? In the rest of Section 4 as well as part of Section 5, we review visual-world studies that explored effects of information-structural and syntactic factors. In the rest of Section 5, we discuss recent work on how semantic factors impact reference resolution, and what this tells us about prominence.

4.1 Pronoun resolution and givenness/topicality

It is widely agreed that factors such as being old/given information and being realized in subject position render an entity a good antecedent for a subsequent pronoun (Brennan, Friedman, & Pollard, 1987; Chafe, 1976; Crawley & Stevenson, 1990, Gordon, Grosz, & Gilliom, 1993; Strube & Hahn, 1996; 1999). These patterns have been observed in corpus studies and a range of different psycholinguistic experiments. Interestingly, many of these factors are also regarded as being connected to the notion of topicality. In fact, the topic of a sentence – the entity that the sentence is about (e.g. Reinhart, 1982; Lambrecht, 1994) – is often regarded as the most prominent referent in the sentence. Broadly speaking, these findings suggest that pronouns prefer referents that are given, topical information.

In this section, we review studies conducted with the visual-world paradigm that have explored how topicality-related factors guide real-time reference resolution. The results show that comprehenders make rapid use of these cues (Arnold, Eisenband, Brown-Schmidt & Trueswell 2000), but also highlight the benefit of disentangling different topicality-related cues (Kaiser, 2011a). The visual-world paradigm is well-suited for investigating how different factors influence referents' prominence, because it allows us to see what potential antecedents comprehenders consider as they hear a pronoun. With eye-movements we can see not only *how strongly* a particular referent is preferred; we can also see *when* this preference begins to emerge.

This can give us more fine-grained information about the strength of different factors, as one might expect a more influential factor to have a stronger effect or, perhaps, for its effect to emerge earlier during processing.

In one of the first visual-world studies to look directly at pronoun resolution, Arnold et al. (2000) investigated how gender information and grammatical role/order-of-mention guide the interpretation of pronouns. Participants listened to sentences like (7) while viewing a scene depicting the two mentioned referents. The task was to indicate whether the sentence matches the picture. In the pictures, either the first- or second-mentioned referent (the subject or the object) matches the description (e.g. either Donald or {Mickey/Minnie} is holding an umbrella).

(7) *Donald is bringing some mail to {Mickey/Minnie} while a violent storm is beginning.*
He's/She's carrying an umbrella, and it looks like they're both going to need it.

When the two referents had different genders, participants' eye-movements revealed a rapid sensitivity to gender cues: People rapidly converged on the intended antecedent, regardless of whether it was the subject or the object of the first sentence. When the two characters had the same gender (e.g. Donald and Mickey), participants' eye-movements reveal a preference for the subject of the first sentence, which fits with the idea that referents in subject position are highly prominent. This preference was strengthened when the prominence of the subject was further boosted by mentioning the subject referent again in an additional clause before the critical sentence (see ex.8).

(8) *Donald is bringing some mail to {Mickey/Minnie}.*
[He's sauntering down the hill,]additional clause while a violent storm is beginning.
He's/She's carrying an umbrella, and it looks like they're both going to need it.

As a whole, these studies show rapid effects of gender information as well as rapid effects of subjecthood/topicality, emerging within the first 200ms after pronoun offset. In a follow-up study using the same items with 4-5-yr-old children, Arnold, Brown-Schmidt & Trueswell (2007) found that while children's eye-movements reveal early effects of gender, children did not seem to be sensitive to subjecthood: In the same-gender conditions, children did

not converge on the target until after it had been lexically disambiguated (e.g. by mention of the umbrella). Arnold et al. suggest that this may be due to children initially preferring more reliable cues that do not depend on skills to manage joint attention. Somewhat divergent results arise in a preferential-looking study by Song & Fisher (2005). Their results suggests that as long as there are enough convergent cues (in their study, subjecthood and repeated mention), then children as young as three start to converge on the intended referent, though not until 1000ms after the pronoun. Furthermore, recent work by Hartshorne et al. (in press) using the visual-world paradigm suggests that over the course of development, children become faster at using subjecthood as a cue. They further note that these effects may have been too slow (occurred too late) to be captured in some of the earlier studies.

Let us now consider why pronoun interpretation is influenced by subjecthood. What is it about subjects that makes them good antecedents? In particular, can the effects of subjecthood be attributed to the fact that subjects are often given information? Arnold et al. (2000) found that pronoun interpretation is guided by subjecthood even when both the subject and the object are new information. This suggests that subjecthood does not need to coincide with givenness in order to influence pronoun interpretation. In related work, Kaiser (2011a) investigated what happens when subjecthood is explicitly pitted against givenness. Can we see effects of subjecthood when it is purposefully dissociated from information-structural considerations?

In Kaiser's study, participants heard mini-stories (ex.9) while looking at scenes showing the mentioned characters. In the critical sentence, either the subject, the object or neither was given information (and pronominalized), resulting in three conditions as shown in (9). The test sentence had a gender-ambiguous pronoun in subject position.

(9)

	Lead-in + critical sentence
Greg is always very supportive of others. **_He_** *congratulated* **_John_** *enthusiastically yesterday.*	[S=pro, O=name] condition
Mike did very well in last month's tennis tournament. **_John_** *congratulated* **_him_** *enthusiastically yesterday.*	[S=name, O=pro] condition

Greg _congratulated **_John_** enthusiastically yesterday._	[S=name, O=pro] condition
The prizes for the best-ranked tennis players were about to be announced, and	Look-away sentence
_he_ _was holding a new yellow tennis racket._	**_Test sentence_**
Everyone was in a good mood that day.	Wrap-up

Participants were instructed that if there was a mismatch between the narrative and the picture, they should use the mouse to click on the region with the error. (In all target items, the test sentence was incorrect with respect to both potential referents. For example, in ex.(9), both men were holding tennis rackets, but neither racket was yellow. Thus, participants' clicks revealed which character they interpreted as the referent of the pronoun.) These off-line click responses revealed a subject preference in all conditions, but this preference was weakened when the object was pronominalized and strengthened when the subject was pronominalized. The eye-movements showed even more clearly the consequences of having a pronominalized, discourse-old object: When the object in the critical sentence was pronominalized, the pronoun in the test sentence triggered consideration of both the pronominalized object and the full-name subject. This contrasts with the other two conditions, which showed a clear subject preference. Thus, these results show that both subjecthood and givenness (strengthened here with pronominalization) influence pronoun interpretation during real-time processing, even when they do not coincide. Importantly, the separable effects of subjecthood and givenness show that subjects have a special status even when they are new information and do not fit the standard definition of 'topic.' Thus, effects of syntactic prominence exist, even when separated from information-structural considerations.

4.2 Comparing effects of topicality and focusing on pronouns

The finding that factors such as subjecthood and pronominalization – often connected to topicality – influence pronoun resolution forms an interesting counterpart to another line of

research suggesting that entities which are contrastively focused also occupy a privileged cognitive status (e.g. Hornby, 1974; Zimmer & Engelkamp, 1981; Sturt, Sanford, Stewart & Dawydiak, 2004; Ward & Sturt, 2007; see also Cutler & Fodor, 1979 for related work). For example, Hornby found that when people were presented with cleft sentences (e.g. "It is the girl who is riding the bicycle"), they were better at detecting mismatches when the mismatching information was focused (the girl) than when it was presupposed (the bicycle). Looking more specifically at reference resolution, Almor (1999) found that NPs referring to clefted antecedents are read faster (i.e., processed more easily) than NPs that refer to non-clefted/non-focused antecedents. Furthermore, on the basis of pronoun resolution patterns, Birch, Albrecht and Myers (2000) argued that focused concepts are more salient and have a 'stronger memory trace' (Birch et al. 2000:302) than non-focused concepts. In more recent work, Foraker and McElree (2006) suggested that clefting makes referents 'more distinctive' in memory.

In order to see how focusing and topicality interact during the process of pronoun resolution, two non-eye-tracking studies by Arnold (1999) and Cowles (2003) (published as Cowles, Walenski & Kluender, 2007) tested what happens when these two notions are pitted against each other. In one experiment, Arnold (1999) used an open-ended story-continuation task to compare given, pronominalized referents in subject position and contrastively focused entities in object position. She tested both SVO sentences with focused objects and object clefts (ex(10)).[8] The results showed that participants were more likely to use pronouns for the preceding subject (Ron in ex.10) than the preceding object (Fred in ex.10), regardless of whether the focused object was clefted or its normal post-verbal position. This suggests that given information in subject position is more prominent, at least from the perspective of pronoun production, than a contrastively focused object. Recent work by Colonna, Schimke & Hemforth (2010) on clefting in French also found a preference for topic over focus.

(10)   a. *Ron was looking through his address book, trying to make up his mind.*

    b. *He had an extra ticket to the opera, but he didn't know which friend to invite.*

        c. *The one he decided on at last was Fred.*   (CLEFTED) or

        c'. *At last he decided on Fred.*   (SVO)

    d. *participant provides a continuation sentence*

---

[8] Arnold used *the one*-clefts, a type of cleft referred to by Prince (1978) as a "cleft with a lexical head."

However, divergent results emerge from Cowles (2003) and Cowles et al. (2007). Cowles used a naming task to investigate three kinds of antecedents: (i) contrastively focused names in subject position, (ii) given-information names in subject position, and (iii) new-information names in subject position (ex.11a-c).

(11a) *Contrastively focused subject:*
A new movie opened in town. It was Anne who called Sarah.

(11b) *Given subject*:
Anne wanted to see the new movie with Sarah. So, Anne called Sarah.

(11c) *New subject:*
 A new movie opened in town. So, Anne called Sarah.

(11d) *Pronoun-containing third sentence (same in all conditions):*
But later that night, she couldn't go to the movie after all.

Naming latencies for visual words presented right after the critical pronoun in the third sentence were numerically faster when participants were asked to name the preceding subject (e.g. Anne) than the preceding object (e.g. Sarah), regardless of whether the subjects were in focus, discourse-old or discourse-new. In contrast to Arnold who found that pronominalized, given subjects are more prominent than contrastively focused objects, Cowles concludes that "two information structure types that are considered distinct….appear to have the same psychological effect" (Cowles 2003:94). The divergent conclusions may be attributable to effects of grammatical role: Arnold's foci were in object position and Cowles' foci were in subject position. Thus, it is not clear whether the interpretation of pronouns – which is argued to be sensitivity to the general notion of prominence – is more sensitive to prominence defined in terms of givenness and topicality or prominence defined as contrastive information.

4.3 Visual-world studies on the effects of topic and focus

To shed light on the divergent claims regarding the effects of topicality and contrastive focus on pronoun interpretation, Kaiser (2011a) tested focused and given entities in both subject and object position. In a visual-world eye-tracking study, participants listened to dialogs like ex.(12) while viewing clip-art scenes.

(12) *Speaker A*: I heard that Greg congratulated Mike enthusiastically yesterday.

Speaker B:

(a) *No, that's not quite right.*

(b)   (i) *He congratulated **John**.*   [SVO.Object=focus]

            (ii) ***John** congratulated him.*   [SVO.Subject=focus]

            (iii) *It was **John** that he congratulated.*   [Cleft.Object=focus]

            (iv) *It was **John** who congratulated him.*   [Cleft.Subject=focus]

(c) *The prizes for the best-ranked tennis players were about to be announced, and*

(d) ***he** was holding a new yellow tennis racket*   [TEST SENTENCE]

(e) *Everyone was in a good mood that day.*

The first sentence of each dialog – uttered by speaker A – was corrected by speaker B. Either the identity of the subject or the object was corrected, using an SVO sentence or an *it-*cleft. Both SVO sentences and *it*-clefts were included to see if potential focus effects are stronger when the focus status of an entity is indicated not only by the discourse context but also by a special syntactic construction. Similar to the experiment on givenness and subjecthood described in Section 4.1, participants were engaged in a picture-verification task.

On the whole, the main pattern that emerged from participants' off-line picture-verification responses is an overarching subject preference in all conditions, regardless of whether the subject is pronominalized or contrastively focused. Eye-movement patterns triggered by the pronoun in the test sentence also showed a subject preference, regardless of whether the subject or object was contrastively focused or pronominalized.

The finding that a subject preference arises even in sentences where the subject is focused and the object is given and pronominalized shows that contrastive focus does indeed boost a referent's prominence. Recall that in a similar configuration, when the subject was new

information (but crucially not contrastively focused) and the object was pronominalized, Kaiser (2011a) found persistent competition between the subject and the object (Section 4.1 above). Now, we find that when the discourse-new subject is contrastively focused (i.e., has two things in its favor: subjecthood and contrastive focus), it is preferred over the given object. These findings show that when grammatical role is taken into consideration, we can detect effects of contrastive focus boosting referents' prominence even in the presence of a discourse-old, pronominalized ('topical') competitor.

In sum, if we treat pronoun resolution as a measure of prominence, then two phenomena which are often regarded as complementary, topic and focus, both have the effect of increasing prominence. This finding, in light of the fact that topic and focus nevertheless differ in many informational-structural and pragmatic respects (e.g., Gundel & Fretheim, 2004), points towards a conceptualization of prominence as a phenomenon that emerges from a wealth of diverse ingredients.

5. Going beyond pronouns: Interpretation of other referring expressions

Much of the psycholinguistic research investigating the effects of prominence on reference resolution has focused on how people interpret ambiguous pronouns. However, languages also use other linguistic forms to refer back to already-mentioned entities, including demonstratives such as *this* and *that*) (e.g. Ariel, 1990; Gundel, Hedberg & Zacharski, 1993). As mentioned above, it is commonly assumed that there exists a correlation between the type of referential form used to refer to an entity and the level of prominence/accessibility of the entity: The more reduced an anaphoric expression is, the more prominent its referent needs to be in the speaker's and listener's mental models of the discourse (e.g. Arnold, 1998:4). Part of the standard accessibility hierarchy is shown in (13). Forms further to the left are used to refer to more accessible/prominent referents:

(13) *null forms > pronouns > demonstratives > full nouns…*

Positing a connection between prominent referents and reduced referring expressions seems plausible. As noted by Garnham, "[a]n expression that has little semantic content … can contribute little or nothing to the identification process, and can only be used where

identification of the referent is either straightforward or not an issue" (Garnham, 2001:55). However, not all referring expression differ in their informativeness. In English, for example. *it* and *that* "are indistinguishable with respect to the description they provide for the intended referent (an inanimate object)" (Ariel, 2001:29), but according to the hierarchy in (13), demonstratives are nevertheless used for less prominent referents than pronouns.

This view of the mapping between referential forms and prominence has been challenged in recent work by Kaiser (2003, 2005) and Kaiser & Trueswell (2008) on Finnish, Brown-Schmidt, Byron & Tanenhaus (2005) on English, and Wilson (2009) on German. Kaiser investigated the interpretation of pronouns and reflexives in Finnish, a language with flexible word order (Section 3.1). In Finnish, human referents can be referred to with both the gender-neutral pronoun *hän* 'he/she' and the demonstrative *tämä* 'this'. Kaiser (2003, 2005) and Kaiser & Trueswell (2008) investigated the interpretation of Finnish pronouns and demonstratives after canonical SVO sentences and noncanonical OVS sentences in a series of experiments, including a visual-world study. To ensure contextual felicity of the SVO and OVS sentences, the preverbal argument was given information (i.e. had been mentioned before) whereas the post-verbal argument was new information. Participants' eye-movement patterns revealed that pronouns and demonstratives are asymmetrical in their referential preferences: Whereas the pronoun *hän* was interpreted as referring to the preceding subject regardless of word order, the demonstrative *tämä* tended to be interpreted as referring to the postverbal argument, especially when it was an object. Thus, Kaiser & Trueswell concluded that the interpretation of the pronoun *hän* is driven primarily by the syntactic role of potential antecedents, while the demonstrative *tämä* exhibits a sensitivity to both word order/information structure and syntactic role. Wilson (2009) found that German pronouns and demonstratives exhibit a similar asymmetry, in that the interpretation of pronouns is more sensitive to grammatical role information whereas demonstratives are more sensitive to discourse-level information (see also Ellert, Järvikivi & Roberts 2009 for visual-world data showing that both German pronouns and demonstratives are sensitive to word order).

Based on their findings, Kaiser & Trueswell argue for a *form-specific, multiple-constraint approach*, where referential forms can differ in how sensitive they are to different antecedent properties. This approach resembles existing multiple-constraint approaches to reference resolution (e.g., Arnold, 1998; Ariel, 1990) in assuming that anaphor resolution is not determined by one constraint but rather is the result of the interaction of multiple constraints. Furthermore,

this approach allows for the multiple constraints that play a role in the interpretation of referential forms to not necessarily carry the same weight for all referential forms.

Related eye-tracking work on Finnish by Järvikivi, van Gompel, Hyönä and Bertram (2005) on the interpretation of the pronoun *hän* when preceded by SVO and OVS sentences found that *hän* exhibits an initial subject preference, followed by later effects of both syntactic role and word order. It is worth noting that the visual-world paradigm is ideal for observing these kinds temporal effects, because it provides an on-going measure of how participants' eye-movement patterns change over time. Although these results differ somewhat from Kaiser & Trueswell (2008) – perhaps due to Järvikivi et al. not using a discourse context before the critical SVO/OVS sentences – they are not incompatible with the main claims of the form-specific approach, namely that the pronoun *hän* and the demonstrative *tämä* differ in how sensitive they are to different types of information.

Further support for the form-specific view comes from an eye-tracking study by Brown-Schmidt, Byron and Tanenhaus (2004, 2005) on *it* and *that* in English. They found that both *it* and *that* are sensitive to extra-linguistic information, such as how easily two objects could be viewed as a composite. For instance, given a command like 'Put the cup on the saucer. Now put that over by the shovel', participants interpreted 'that' as referring to the composite 'cup-and-saucer' 88% of the time. Crucially, Brown-Schmidt et al's results as a whole did not fit with the claim that that *it* is used for more prominent antecedents than *that* (see the hierarchy in (13)), because *it* and *that* did not exhibit the predicted prominence differences.

Additional cross-linguistic evidence for the form-specific approach comes from a visual-world experiment and additional off-line studies and Kaiser (2011b) on Dutch emphatic 'strong' pronouns, non-emphatic 'weak' pronouns and demonstratives. The results indicate that while the differences between demonstratives and nonemphatic personal pronouns correlate with the antecedent's grammatical role, the distinction between strong and weak pronouns cannot be satisfactorily explained by grammatical role, contrary to what some accounts predict. The finding that even the fine-grained, time-sensitive data from visual-world eye-tracking shows no effects of a grammatical-role effect for strong vs. weak pronouns provides convincing evidence that the distinction between these forms is not conditioned on this dimension. If the method had been less sensitive or off-line, one could have argued that perhaps there were weak or transient effects of grammatical role that we simply failed to pick up. However, the visual-world paradigm suggests

that this is not the case. Instead, it appears that the strong form is sensitive to the presence of contrast, showing that referring expressions can differ in what kind of information they are most sensitive to. Form-specific behavior has also been observed with intra-sentential anaphora: Kaiser, Runner, Sussman & Tanenhaus (2009) used a visual-world paradigm to investigate the reference resolution in picture-NP constructions (e.g. *John told Peter about the picture of him/himself on the wall*), and found that pronouns are relatively more sensitive to pragmatic/semantic factors than reflexives.

As a whole, these visual-world studies suggest that the process of reference resolution, once we assume a broader view that extends beyond pronouns, may not be straightforwardly reducible to a unified notion of prominence or salience.

5.1 Another take on prominence: Semantic effects

Recently, the visual-world paradigm has been used to investigate how semantic factors influence reference resolution, including verb semantics and the semantic relations between sentences. For example, Pyykkönen & Järvikivi (2010) tested the interpretation of pronouns after implicit causality verbs such as *fear* and *frighten* in Finnish. The class of implicit causality verbs, especially when followed by the connective *because*, is well-known for influencing pronoun interpretation. For example, 'Mark feared Andy because he…' tends to elicit continuations about Andy (the preceding object), whereas 'Mark frightened Andy because he…' tends to elicit continuations about Mark (the subject). These patterns have been observed repeatedly with a range of methodologies (e.g. Garvey & Caramazza, 1974; Koornneef & Van Berkum, 2006). In their experiment, Pyykkönen & Järvikivi (2010) used the visual-world paradigm to investigate both object-oriented and subject-oriented implicit causality verbs in Finnish. Crucially, participants' eye-movements revealed effects of the verbs' implicit causality biases even before participants encountered the causal connective 'because' (and thus also before the pronoun at the start of the next clause). This finding – clearly revealed thanks to the visual-world paradigm – is important because it shows that implicit causality effects kick in early, and suggests that they are not triggered by the presence of the pronoun but evoked by the verb itself.

These findings connect with another, related line of research investigating how the semantic relations between sentences influence reference resolution. As noted by Kehler (2002) and Kehler, Kertz, Rohde & Elman (2008), pronoun interpretation is influenced by the coherence

relation between the pronoun-containing clause and the preceding clause. In ex.(14), a subject interpretation of 'him' is more likely with a cause-effect relation (14a) than with a similarity relation (ex.14b). (See also Venditti et al. 2001, Tavano & Kaiser 2008 for eye-tracking work with stressed pronouns and different coherence relations). In fact, according to Kehler, pronoun resolution is largely a by-product of general inferencing about inter-clausal relations.

(14) *Phil tickled Stan, and Liz poked him.*
(a) *Phil tickled Stan, and as a result Liz poked him$_{Phil}$*
(b) *Phil tickled Stan, and similarly Liz poked him$_{Stan}$*

Recent work by Kaiser (2009) used visual-world eye-tracking combined with priming to investigate the nature of these coherence relations. In one experiment, participants were presented with visuo-spatial primes, silent video clips that encoded (i) cause-effect/result relations, (ii) similarity relations or (iii) other/neutral relations (e.g., in a Result video prime, a triangle knocks into a circle which falls off a ledge). In another experiment, the coherence relation primes were linguistic (e.g. participants read "The patient pressed the red emergency button near the bed and a nurse quickly ran into the room" for Result). Participants were then shown a target scene with three characters and heard a sentence with an ambiguous object-position pronoun, e.g. *Phil linded Stanley and Kate hepped him.* (Nonce words were used to eliminate effects of verb semantics). The task was to use the mouse to click on the last-mentioned referent (i.e., the antecedent of *him*). Participants' eye-movements to the preceding subject vs. object after different kinds of primes showed that pronoun interpretation can be primed by coherence relations in preceding linguistic and visual input – even when primes and targets are connected only on the level of abstract coherence relations, and when primes are presented in a non-linguistic modality. This points to the existence of shared representations between coherence-related inferencing and pronoun resolution processes.

As a whole, these findings suggest that if we want to conceptualize reference resolution as a process that is sensitive to how prominently different entities are represented in people's minds, our view of what factors influence referents' prominence needs to be expanded to include not only entity-related factors such as grammatical position or givenness, but also event-related factors such as verb semantics and the connections between different events in the discourse (e.g.

result vs. similarity). The visual world paradigm is very well-suited for helping us to better understand the interplay of these factors, because (i) the nature of the paradigm makes it possible to manipulate a variety of contextual features and because (ii) people's eye-movement patterns allow us to track, moment by moment, what aspects of the visual scene they are directing their attention towards and how their attention shifts over time – which in turn gives us a window into the real-time comprehension process.

7. Conclusions

Visual-world eye-tracking studies have made important contributions to our understanding of language processing on the discourse level. As we saw in this chapter, the visual-world paradigm has allowed researchers to explore many aspects of discourse-level processing with great success, including gaining new insights into how comprehenders use prosodic, syntactic and semantic cues during real-time comprehension. As a whole, the results highlight the rapidity with which the human language processing system is capable of making use of discourse-level information, whether it be encoded in pitch accents, word order or the form of referring expressions. This kind of information, which the visual-world paradigm is well-suited for measuring, suggests that discourse-level comprehension should not be relegated to a secondary stage of processing and occurs in tandem with other aspects of language comprehension, such as lexical access and syntactic processing. More generally, the visual-world paradigm has been used to show that processing is not as 'staged' as some other methods might have led us to believe (see e.g. Tanenhaus et al. (1995)'s seminal study, the first to use visual-world eye-tracking for investigating real-time language processing, which showed that even during the earliest moments of processing, visual context influences word recognition and syntactic processing.)

The notion of 'prominence' has played a central role in many of these investigations, in the shape of prosodic prominence (associated with new information), syntactic prominence (often associated with given or topical information) and representational prominence/accessibility in the domain of reference resolution. We've seen that comprehenders use prominence-related information to guide discourse-level aspects of processing, but that the prosodic prominence and syntactic prominence have different information-structural correlates. Furthermore, we also saw that – if we want to conceptualize reference resolution as a process

that is sensitive to how prominently different entities are represented in people's minds – then our view of what factors influence referents' prominence needs to include not only entity-related factors (e.g. givenness), but also event-related factors such as verb semantics and the coherence relations between different events in the discourse.

References

Allopenna, P. D, Magnuson, J.S. & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition: evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419-439.

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis, *Psychological Review*, 106(4), 748-765.

Altmann, G.T.M. & Kamide, Y. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73.247–64.

Altmann, G.T.M. (2004) Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition* 93, 79–87

Ariel, M. (1990). *Accessing NP antecedents*. London: Routledge, Croom Helm.

Ariel, M. (2001). Accessibility Theory: an overview. In T. Sanders, J. Schilperoord & W. Spooren (Eds.), Text Representation, Linguistic and psycholinguistic aspects (pp. 29-87). Amsterdam/Philadelphia: John Benjamins.

Arnold, J. E. & Tanenhaus, M. K. (2011). Disfluency effects in comprehension: how new information can become accessible. In Gibson, E., & Perlmutter, N. (Eds) *The processing and acquisition of reference,* MIT Press.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S, and Trueswell, J. C. (2000). The immediate use of gender information: eyetracking evidence of the time-course of pronoun resolution. *Cognition* 76, B13-B26.

Arnold, J. E., Wasow, T., Losongco, A., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76, 28-55

Arnold, J.E., Brown-Schmidt, S. & Trueswell, J.C. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes,* 22(4), 527-565.

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes,* 11, 17–67.

Beckman, M.E., & Ayers, G M. (1997). *Guidelines for ToBI labelling*, vers 3.0. Ohio State University.

Birch, S. L., Albrecht, J. E., & Myers, J. L. (2000). Syntactic focusing structures influence discourse processing. *Discourse Processes*, 30, 285–304.

Birner, B. & Ward, G. (1998). *Information Status and Noncanonical Word Order in English.* Amsterdam/Philadelphia: John Benjamins

Birner, B. & Ward, G. (2009). Information Structure and Syntactic Structure. *Language and Linguistics Compass* 3/4, 1167–1187.

Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford: Stanford University Press.

Brennan, S. E., Friedman, M. A., & Pollard, C. J. (1987). A centering approach to pronouns. In *Proceedings of the 25th annual meeting of the Association for Computational* Linguistics (pp. 155- 162). Stanford, CA: Association for Computational Linguistics.

Brown-Schmidt, S. (2005) *Language Processing in Conversation.* Ph.D. Dissertation, University of Rochester.

Brown-Schmidt, S., Byron, D.K, & Tanenhaus, M. (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language*, 53, 292-313.

Büring, D. (1997). *The meaning of topic and focus—The 59th Street Bridge accent.* London: Routledge.

Chafe, W.L. (1974). Language and Consciousness. *Language* 50:111– 133.

Chafe, W.L. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View. In: C. Li (ed.). *Subject and Topic*. New York: Academic Press, 25- 55

Chen, A., Den Os, E., & De Ruiter, J. P. (2007). Pitch accent type matters for online processing of information status: Evidence from natural and synthetic speech. *The Linguistic Review,* 24(2), 317-344.

Chomsky, Noam. (1971). Deep structure, surface structure, and semantic interpretation. In D. Steinberg and L. Jacobovits (eds.),*Semantics*. Cambridge: Cambridge University Press. 183-216.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language*. New York: Harcourt Brace Jovanovich.

Clark, H. H., & Haviland, S. (1977). Comprehension and the given-new contract. In R. Freedle (Ed.), *Discourse production and comprehension* (pp. 1-40). Hillsdale, NJ: Erlbaum.

Clifton, Charles, & Lyn Frazier (2004). Should given information come before new? Yes and no. *Memory and Cognition*, 32(6):886–895.

Colonna, S., Schimke, S. & Hemforth, B. (2010). Le rôle de la structure informationnelle dans l'interprétation d'une anaphore pronominale inter-phrastique. In F. Neveu at al. (Eds.), *Congrès Mondial de Linguistique Française,* 1489-1499.

Cooper, R.M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology* 6, 84–107.

Cowles, H.W. (2003). *Processing Information Structure: Evidence from Comprehension and Production.* Ph.D. dissertation, UCSD.

Cowles, H.W., Walenski, M. & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: Topic, constrastive focus and pronouns. *Topoi,* 26, 3-18.

Crawley, R. J., & Stevenson, R. J. (1990). Reference in single sentences and in texts. *Journal of Psycholinguistic Research*, 19(3), 191-210.

Cutler, A., & Fodor, J. (1979). Semantic focus and sentence comprehension. *Cognition,* 7, 49-59

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language, 47*, 292–314.

Delin, J. & Oberlander, J (1995). Syntactic constraints on discourse structure: the case of it-clefts. *Linguistics*, 33, 3.

Dufter, A. (2009). Clefting and discourse organization: Comparing Germanic and Romance. In A. Dufter & D. Jacob (eds.), *Focus and Background in Romance Languages (Studies in Language Companion Series 112).* Amsterdam: John Benjamins.

Ellert, M., Järvikivi, J. & Roberts, L. (2009) Information structure guides gaze behavior: processing the German subject pronouns *er* and *der* in spoken discourse. Poster presented

at 15th *Annual Conference on Architectures and Mechanisms for Language Processing*, Barcelona, Spain.

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology, 63,* 639-645.

Féry, C, Skopeteas, S. & Hörnig, R. (2010). Cross-linguistic comparison of prosody, syntax and information structure in a production experiment on localizing expressions. *Transactions of the Philological Society*, 108:3. 329–351

Foraker, S. & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language,* 56(3), 357-383

Garnham, A. (2001). *Mental models and the interpretation of anaphora.* Hove: Psychology Press.

Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic Inquiry*, 5, 459-464.

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study.* Amsterdam: John Benjamins.

Gleitman, L., January, D., Nappa, R. & Trueswell, J. (2007). On the give and take between event apprehension and utterance formulation. Journal of Memory and Language, 57, 544-569.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering o attention in discourse. *Cognitive Science*, 17, 311-347.

Green, M. & Jaggar, P. (2003). 'Ex-situ and In-situ Focus in Hausa: syntax, semantics and discourse', in Lecarme, J (Ed.), *Research in Afroasiatic Grammar II.* [CILT 241]. Amsterdam: John Benjamins: 187-213.

Griffin, Z. M., & Bock, J. K. (2000). What the eyes say about speaking. Psychological Science, 11, 274-279.

Gundel, J. K. (1974). *The role of topic and comment in linguistic theory*. Ph.D. dissertation, University of Texas at Austin.

Gundel, J. K. (1988). Universals of topic-comment structure In M. Hammond, E. Moravczik and J. Wirth (eds.), Studies in syntactic typology, 209-239, Amsterdam: John Benjamins, 209-239.

Gundel, J.K. & Fretheim, T. (2004). Topic and Focus. In G. Ward & L. Horn (Eds.), *Handbook of Pragmatics (Blackwell Handbooks in Linguistics)*. (pp.175-196). Oxford: Blackwell.

Gundel, J.K., Hedberg, N. & Zacharski, R. (1993). Cognitive status and the form of referring Expressions in Discourse. *Language,* 69, 274-307.

Gussenhoven, C. (1983). Focus, mode, and nucleus. *Journal of Linguistics, 19,* 377–417.

Halliday, M. (1967). Notes on Transitivity and Theme in English. Part 1 and 2. *Journal of Linguistics* 3, 37-81; 199-244.

Hartshorne, Joshua K., Rebecca Nappa, & Jesse Snedeker (in press). Development of the First-Mention Bias. *Journal of Child Language*.

Haviland, S.E., & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13: 512-521.

Hedberg, N. (1990). *Discourse Pragmatics and Cleft Sentences in English*. Ph.D dissertation, Universitiy of Minnesota.

Hedberg, N. (2000). The referential status of clefts. *Language* 76: 891-920.

Hornby, P.A. (1974). Surface structure and presupposition. *Journal of Verbal Learning and Verbal Behavior,* 13, 530-538.

Huang, Y. & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. Cognitive Psychology, 58(3), 376-415.

Hwang, Heeju & Kaiser, Elsi. (2014). The role of the verb in grammatical function assignment in English and Korean. *Journal of Experimental Psychology: Learning, Memory, and Cognitio*n, 40: 1363-1376.

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*.Cambridge, MA: MIT Press

Järvikivi, J., Van Gompel, R.P.G., Bertram, R., & Hyönä, J. (2005). Ambiguous pronoun resolution: Contrasting the first-mention and subject preference accounts. *Psychological Science*, 16, 260-264.

Kaiser, E. & Trueswell, J.C. (2008) Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, 23(5), 709-748.

Kaiser, E. & Trueswell, J.C. (2004) The role of discourse context in the processing of a flexible word-order language. *Cognition* 94(2), 113-147.

Kaiser, E. (2009). Effects of Anaphoric Dependencies and Semantic Representations on Pronoun Interpretation. In S.L. Devi, A. Branco and R. Mitkov (eds). *Anaphora Processing and Applications*, pp.121-130. Heidelberg: Springer.

Kaiser, E., Runner, J.T., Sussman, R.S. & Tanenhaus. M.K. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition*, 112, 55-80.

Kaiser, Elsi (2011a). Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus, *Language and Cognitive Processes*, 26, 1625-1666.

Kaiser, Elsi. (2011b). Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives, *Language and Cognitive Processes*, 26, 1587-1624.

Kamide, Y & Altmann, G.T.M. & Haywood, S. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye-movements. *Journal of Memory and Language,* 49, 133–59.

Kehler, A. (2002). Coherence, reference, and the theory of grammar. CSLI Publications, Stanford.

Kehler, A., Kertz, L., Rohde, H., Elman, J. (2008). Coherence and Coreference Revisited. *Journal of Semantics (Special Issue on Processing Meaning),* 25(1), 1-44.

Kiss, K. E. (1998). Identificational focus versus information focus. *Language*, 74, 245-273.

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.

Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, 54, 445-465.

Ladd, D.R. (1996). *Intonational Phonology.* Cambridge: Cambridge University Press.

Lambrecht, K. (1994.) *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents.* Cambridge: Cambridge University Press.

Lambrecht, K. (2001). A framework for the analysis of cleft constructions. *Linguistics,* 39, 463-516.

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, 31, 133-156.

Myhill, J. (1992). *Typological discourse analysis*. Oxford: Basil Blackwell.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.

Prince, E.F. (1978). A comparison of WH-clefts and IT-clefts in discourse. *Language,* 54, 883-906.

Prince, E.F. (1992). The ZPG letter: subjects, definiteness, and information status. In S. Thompson and W. Mann (Eds.), *Discourse description: diverse analyses of a fund-raising text* (pp. 295–325). Amsterdam/Philadephia: John Benjamins.

Pyykkönen, P. & Järvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Experimental Psychology*, 57 (1), 5-16.

Pyykkönen, P., Matthews, D., & Järvikivi, J. (2010). Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes,* 25, 115-129.

Reinhart, T. (1982). *Pragmatics and linguistics: an analysis of sentence topics*. University of Indiana Linguistics Club. (also *Philosophica* 1981, 27, 53-94).

Rochemont, M. (1986). *Focus in Generative Grammar.* Amsterdam: Benjamins

Rooth, M. (1992). A Theory of Focus Interpretation. *Natural Language Semantics,* 1, 75-116.

Schwarzschild, R. (1999). GIVENness, Avoid F and other constraints on the placement of focus. *Natural Language Semantics,*7, 141–177.

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

Sekerina, I.E. & Trueswell, J.C. (2012). Interactive processing of contrastive expressions by Russian children. *First Language* 32: 63-87.

Selkirk, E. O. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 550–569). Oxford, UK: Blackwell.

Sgall, P., & Hajicova, W.E. (1977). Focus on focus. *The Prague Bulletin of Mathematical Linguistics,* 28: 5-54.

Song, H., & Fisher, C. (2005). Who's 'she'? Discourse prominence influences preschoolers comprehension of pronouns. *Journal of Memory and Language,* 52, 29–57.

Spivey, M.J., Richardson, D.C., & Fitneva, S.A. (2004). Thinking outside the brain: Spatial indices to visual and linguistic Information. In J. Henderson & F. Ferreira (Eds.), *Interfacing Language, Vision, and Action*. pp 161-190. San Diego, CA: Academic Press.

Steedman, M. (2000). Information structure and the syntax–phonology interface. *Linguistic Inquiry*, 31, 649–689.

Strube, M., & Hahn, U. (1996). Functional centering. In *Proceedings of the 34th annual meeting of the Association for Computational Linguistics* (pp. 270-277), Santa Cruz, CA.

Strube, M., & Hahn, U. (1999). Functional centering: Grounding referential coherence in information structure. *Computational Linguistics,* 25(3), 309-344.

Sturt, P., Sanford, A. J., Stewart, A., & Dawydiak, E. (2004). Linguistic focus and good-enough representations: an application of the change-detection paradigm. *Psychonomic Bulletin and Review*, 11, 882-888.

Tanenhaus, M. K. (2007). Spoken language comprehension: insights from eye movements. In G. Gaskell (Eds.), pps. 309-326. *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press.

Tanenhaus, M.K. Spivey-Knowlton, M., Eberhard, K.M & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 1632–1634.

Tavano, E. & Kaiser, E. (2008). Effects of stress and coherence on pronoun interpretation. Poster presented at the *21st annual CUNY Conference on Human Sentence Processing*, University of North Carolina-Chapel Hill

Vallduví, E. & Vilkuna, M. (1998). On rheme and kontrast. In Culicover, Peter and Louise McNally (eds.), *The Limits of Syntax. Syntax and Semantics* 29. New York: Academic Press, 79-108.

Vallduvi. E. (1990). *The information component*. Ph.D dissertation, University of Pennsylvania.

Van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing "easy" and "hard" events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, 71(1), 124-144.

Venditti, J. J., Stone, M., Nanda, P. & Tepper, P. (2001). Discourse constraints on the interpretation of nuclear-accented pronouns. In *Proceedings of the 2002 International Conference on Speech Prosody*, Aix-en-Provence, France.

Vilkuna, M. (1989) *Free Word Order in Finnish: Its Syntax and Discourse Functions.* Helsinki: Finnish Literature Society.

Ward, G. (1985). *The semantics and pragmatics of preposing*. Ph.D dissertation, University of Pennsylvania.

Ward, P. & Sturt, P. (2007). Linguistic focus and memory: an eye-movement study. *Memory and Cognition*, 35, 73-86.

Weber, A., Braun, B., & Crocker, M. W. (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech, 49*, 367–392.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eyemovements. *Cognition*, 99, B63–B72.

Wilson. F. (2009). *Processing at the syntax-discourse interface in second language acquisition.* Ph.D. dissertation, University of Edinburgh.

Yee, E., Heller, D., & Sedivy, J.C. (2009). On the relationship between eye-movements and activation: Active vs. passive tasks during ambiguous pronoun resolution. Poster presented at the 22nd annual CUNY Conference on Human Sentence Processing.

Zimmer, H. D. & Engelkamp, J. (1981). The given-new structure of cleft sentences and their influence on picture viewing. *Psychological Research,* 43, 375-389.