Cogn Neurodyn (2014) 8:353–361
DOI 10.1007/s11571-014-9305-1

REVIEW ARTICLE

A review on the cognitive function of information structure during language comprehension

Lin Wang · Xiaoqing Li · Yufang Yang

Received: 30 December 2013 / Revised: 9 July 2014 / Accepted: 22 July 2014 / Published online: 1 August 2014

Abstract During communication, information structure can be used to highlight the most relevant piece of information, so that sufficient amount of attention can be allocated to the most important information. This paper aims to review the cognitive function of information structure during language comprehension from a neurocognitve perspective. First, we gave a brief introduction to the concept of information structure that has been studied mostly in linguistic field. Then we introduced recent studies on information structure using electrophysiological and neuroimaging techniques. After that, we discussed the relationship between attention and language processing more generally. Finally, we discussed potential directions for future studies.

Keywords Information structure · Focus · Attention · Language processing · ERP · fMRI

Introduction

Picture yourself in a party while a stranger talks to you about how to build a computer in which you have no interest at all. Your mind might drift away thinking about your hungry dog at home. In this case, you will not be able to understand him because you simply do not pay enough attention to his words. However, a question such as "what do you think?" or a word with suddenly raised pitch in his sentence might grab your attention and then you start to try to understand what he is talking about. In this scenario, attention plays an important role in how well you are engaged in the conversation.

Specifically, there are two sets of processes that control the flow of information. The top-down control refers to the ability to selectively attend to the relevant piece of information based on prior knowledge or current goals, whereas the bottom-up selection refers to the capacity of quickly allocating processing resources to novel or salient stimuli irrespective of prior goal or task. These two processes have been termed as two aspects of attention, and have been evidenced in numerous studies of visual perception (for a review see Corbetta et al. 2002). The relationship between attention and language processing has been discussed previously. For instance, the visual world paradigm assumes that language-mediated eye movement reflects overt attention guided by language input. Since the language users' eye movements are systematically related to their linguistic processing, the visual world paradigm has been used to study the way listeners understand and speakers produce utterances (for a review see Huettig et al. 2011). Meanwhile, it has been shown that the executive control ability (including attentional control) of bilinguals is more advanced than that of monolinguals (Bialystok 1999; Bialystok and Viswanathan 2009), suggesting the influence of language experience on cognitive development. However, the above-mentioned studies have not addressed how language processing itself is modulated by different attentional settings. In particular, little is known how linguistic device is used to modulate attention, which in turn affects language processing. Therefore, this paper aims to provide a comprehensive insight into the interface between attention and language comprehension by reviewing recent studies on

L. Wang (✉) · X. Li · Y. Yang (✉)
Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Lincui Road 16, Chaoyang District, Beijing 100101, China
e-mail: wangl@psych.ac.cn

Y. Yang
e-mail: yangyf@psych.ac.cn

how linguistic marking—information structure (IS)—modulates language processing from a neurocognitive perspective.

In this paper, we first introduced some concepts regarding IS, then we presented empirical studies on the IS influence of language processing. Among these studies, the role of IS in modulating depth of processing was highlighted. After that, the relationship between attention and language processing was discussed in general. Finally, future directions on the interaction between language and attention were pointed out.

Some concepts of information structure

People make use of multiple cues to organize their utterances and similarly interpret others' utterances in order to make their communication efficient. Take the following conversation as an example: Speaker A asks *What did mum buy to cook?* Speaker B answers *Mum bought* ***beef*** *to cook.* Speaker A brings up a topic and opens a slot for speaker B to fill in. Then speaker B fills in the slot with the relevant information and emphasizes it by putting pitch accent on it (the word in bold). As a result, speaker A rapidly directs his/her attention towards the right position in the sentence and finds the required information. This way of packing different elements of information in sentences is referred to as information structure (IS) (Halliday 1967; Jackendoff 2002; Lambrecht 1996; Erteschik-Shir 2007). It specifies what information in the sentences is important to the interlocutors, so that the listener knows where to target for the relevant information. In this sense, IS can be used to highlight the most relevant piece of information and thus guide the attention of listeners or readers to the relevant information.

Different terminologies have been used to describe the distinctive elements in sentences invoked by IS, such as background-focus, topic-comment, theme-rheme, presupposition-focus, given-new (Chomsky 1965; Karttunen and Stanley 1979; Paterson et al. 2007; Steedman 1991). Among these distinctions, background is generally, although not always, corresponding to topic, theme, presupposition and the given, while focus is corresponding to comment, rheme and the new. Since this paper does not aim at further fine-grained conceptual distinctions, we will use the distinction of background-focus to cover the relevant distinctions in IS status. In this distinction, background refers to information that is shared by the interlocutors, while focus is new or contrastive information that is salient or important for listeners and readers (Gűnther et al. 1999).

There are several ways to mark IS, including question contexts in question–answer pairs (e.g., *what did mum buy to cook?/who bought beef to cook?*), prosodic features such as pitch accent (e.g., *Mum bought* ***beef*** *to cook.* The focus is in boldface), syntactic construction like it-cleft structure (e.g., *It is* ***beef*** *that mum bought to cook.*), word order (e.g., ***Beef****, mum bought to cook.*), focus-marking particles (e.g., *Mum ONLY bought* ***beef*** *to cook.*, the particle is capitalized), and information contrast (e.g., *Mum did not buy pork, she bough* ***beef*** *to cook.*). Also, the way to express IS varies across languages (Gussenhoven 2008). For instance, pitch accent is an important cue for differing between focused and non-focused information in English and Dutch languages, whereas it does not affect listeners' perception or interpretation of speech in Italian language (Swerts et al. 2002). Nevertheless, the function that IS plays during language processing seems to be universal.

The linguistic functions of IS have been widely discussed in the linguistic literature. The discussions have mainly focused on describing multiple linguistic devices in marking IS and the interactions both within these markers and between the IS markers and other language layers (such as phonological, syntactic and semantic aspects). However, this review will focus on the cognitive functions of IS during language comprehension. Given their non-invasive nature, the event-related potential (ERP) and functional magnetic resonance imaging (fMRI) techniques have been extensively used in language studies. In the next section, we will present some empirical studies that have utilized these techniques to investigate the role of IS in language comprehension.

Empirical studies on the IS influence on language processing

This section will introduce two aspects of IS influences on language processing. The first is how different IS markers interact to modulate language comprehension. The other is how IS markers modulate the depth of language comprehension.

Interactions between different IS markers during language comprehension

Some experimental studies have examined how the correspondence between different markers of IS (e.g., prior context vs. pitch accent; prior context vs. syntactically marked focus) influences language processing. Psycholinguistic studies using behavioral measures found that: listeners were more likely to judge sentences as appropriate when the new information was accented and the given information was unaccented (Birch and Clifton 1995); speech processing was facilitated when the new (or focused) information was accented and the given information was de-accented (Bock and Mazzella 1983; Dahan et al. 2002; Terken and Nooteboom 1987). An eye-tracking study also showed that listeners rapidly interpreted an accented word as referring to new referent, and de-accented word as referring to given referent (Dahan et al. 2002). Therefore, listeners are sensitive to the mapping between accentuation and new/given information status.

Behavioral studies have been fruitful in demonstrating the influence of IS on language comprehension. Since there has been an excellent review paper on the behavioral studies of IS processing (see Cowles 2012), this paper focused on the neurological aspect of IS processing. In recent years, increasing number of ERP studies have investigated the temporal aspects of the influence of IS on language comprehension. These studies mainly manipulated the correspondence between pitch accent (in auditory sentences) and information status of a word that was marked in different manners. For missing accentuation (i.e., focused information is unaccented), most studies found negative shifts compared to accented conditions, with varying time windows and scalp topographies: an anterior negative effect between 100 and 500 ms in English (Johnson et al. 2003); an anterior negative effect between 250 and 1,500 ms in Japanese (Ito and Garnsey 2004); a centro-parietal N400 effect between 200 and 600 ms in German (Hruska and Alter 2004); a broadly distributed negative effect between 250 and 450 ms (Bögels et al. 2011) or between 200 and 500 ms (Dimitrova et al. 2012) in Dutch, a sustained central posterior negative deflection lasting about 500 ms in German (Toepel et al. 2007); broadly distributed negative effects between 150 and 1,050 ms for sentence-final words in French (Magne et al. 2005). However, another two studies found a positive shift during 100–750 ms over left hemisphere in German (Heim and Alter 2006, 2007). In addition, a positive effect between 300 and 900 ms in central posterior areas was found in Chinese (Li et al. 2008).

Some attempts have been made to explain the seemingly opposite patterns of ERP responses to pitch accent during language comprehension. Based on previous studies, Li and Yang (2013) proposed that the observed ERP effects in response to missing pitch accent of the new information could be accounted for by two possible mechanisms. First, the larger negativity for the unaccented new information (Johnson et al. 2003; Hruska and Alter 2004; Bögels et al. 2011; Dimitrova et al. 2012) indicates the processing difficulty of the words because pitch accent conveys sentence-level meaning that relates to the focus distribution of the sentence. This was supported by behavioral studies that showed increased comprehension time when the accent placement was inappropriate in the sentences (Bock and Mazzella 1983; Terken and Nooteboom 1987). Second, the larger negativity for the accented new information (or the larger positivity for the unaccented new information; Heim and Alter 2006, 2007; Li et al. 2008) might reflect increased unification load of the words because more attentional resources were allocated to the accented information. In order to test these two assumptions, Li and Yang (2013) used question context (e.g., *What kind of activity does Mingming usually select for relaxation after class?*) to mark critical word in the target sentence as focus (e.g., *He said that Mingming usually selects* ***listening to songs/ cooking*** *for relaxation.* Critical words are in boldface). The critical word conveyed new information and was either highly expected (e.g., *listening to songs*) or lowly expected (e.g., *cooking*) based on the previous context (e.g., *something for relaxation*). In addition, the critical word was either realized with pitch accent or not. They found that the unaccented new information elicited a larger negativity than the accented new information when the words were lowly expected, whereas the opposite effect was found when the words were highly expected. The results suggest that the ERP responses to the inappropriate pitch accent (i.e., new information—not accented) depend on the semantic accessibility of the critical words. When the new information was lowly expected, the information status of the critical word conveyed by the lack of pitch accent (i.e., being given information) mismatched with that indicated by the preceding context (i.e., being new information), thereby causing increased negativities for the unaccented condition. When the new information was highly expected, the critical word might have been pre-activated from the context, so pitch accent could guide the listeners to allocate more attention to the critical word for deeper processing, resulting in larger negativities for the accented condition. In addition, they examined oscillatory brain activities during sentence comprehension. They found that the missing accentuation of the lowly expected new information induced larger theta power increases, whereas the accentuation of the highly expected new information induced larger alpha power decreases. Since the theta power increase has been related to increased lexico-semantic retrieval (Bastiaansen et al. 2005, 2008) and the alpha power decrease has been related to increased attention (Jensen et al. 2012; Shahin et al. 2009), the oscillatory effects further support the view that pitch accent can both convey sentence-level meaning and modulate attention allocation.

ERP responses for the spurious accentuation (i.e., non-focused information is accented) were even more divergent: Li et al. (2008) found an N400 effect in the time window of 200–700 ms over central posterior region, while Li et al. (2011) reported a frontal-central negative effect (270–510 ms). As mentioned above, the negative effect might reflect the processing difficulty of the words due to the mismatch between their information status conveyed by context (being given) and that conveyed by pitch accent (being new). On the contrary, Magne et al. (2005) found a positive effect between 300 and 600 ms in frontal area and Bögels et al. (2011) reported a positive effect in the time window of 250–450 ms with a broad distribution. In addition, Dimitrova et al. (2012) found an early positive shift around 100 ms followed by a right-lateralized negative effect (N400) as well as a late positivity on midline electrodes. The positive shift for the accented compared to unaccented (given) information has been proposed to reflect attentional resources captured by the prominence of accentuation. Furthermore, Heim and Alter (2006) and Heim and Alter (2007) argued that ERP responses to extra accentuation were dependent on the relative position between accentuation and focused information, with an positive shift between 150 and 250 ms when the extra accentuation was far ahead of focused information and a negative shift over central sites between 100 and 550 ms when the extra accentuation was next to the focused information. These findings suggest that the ERP responses to the inappropriate accentuation (i.e., given information—accented) might depend on the predictability of the pitch accent: When the given information was far ahead of focused information, the constraint of given information being unaccented was still weak, so the extra accentuation might have captured attention which led to a positive effect; however, when the given information was next to the focused information, the expectation of placing no accentuation on the given information became strong, so the spurious accentuation caused processing difficulty and thus led to a negative effect. However, no specific effect was found in some other studies (Hruska and Alter 2004; Ito and Garnsey 2004; Johnson et al. 2003; Toepel et al. 2007).

The reported ERP effects suggest that the interaction between pitch accent and information status starts as early as 100 ms and lasts as late as 1,500 ms, indicating that people can rapidly detect the mismatches between various IS markers. However, given the various effects in response to the inappropriately pronounced words, it is difficult to draw any conclusion regarding the functional significance of the effects. The heterogeneity of these results may be partly explained by between-study differences. For instance, these studies differ in languages (German, Dutch, French and Chinese), marking of information status (question–answer pairs: Dimitrova et al. 2012; Heim and Alter 2006; Hruska and Alter 2004; Ito and Garnsey 2004; Johnson et al. 2003; Magne et al. 2005; Toepel et al. 2007; short discourse: Li et al. 2008; picture context: Bögels et al. 2011; focus particle such as "even" in sentence: Heim and Alter 2007), employed tasks (to judge the appropriateness of the prosodic pattern: Heim and Alter 2007; Hruska and Alter 2004; Ito and Garnsey 2004; Johnson et al. 2003; Magne et al. 2005; Toepel et al. 2007; to evaluate the discourse meaning: Bögels et al. 2011; Dimitrova et al. 2012; Heim and Alter 2006; Li et al. 2008), semantic accessibility of the words (Li and Yang 2013), as well as the position of the mismatch between pitch accent and focused information (sentence initial, medial or final).

In addition to studies on the processing of IS during listening, some studies examined the processing of focused information during reading. Cowles et al. (2007) used it-cleft structure to highlight an inquired word in a question context either appropriately or inappropriately (e.g., *Who did the queen silence with a word, the banker or the advisor? It was the* ***banker*** *that the queen silenced./It was the* ***queen*** *that silenced the banker.*). They reported N400 effects in response to the inappropriately highlighted words, suggesting that the readers immediately detected the mismatch between the information state marked by syntactic structure and that marked by discourse context (Cowles et al. 2007). Since the N400 has been related to lexical and semantic processing (for a review see Kutas and Federmeier 2011), the observed N400 effect suggests that the violation of IS increased semantic processing difficulty. They have also found that words in the focus position (regardless of the appropriateness) elicited larger positivities than words in other position of the sentence except for the final word, which could be associated with greater effort of integrating focused information. Combined with the results coming from spoken language processing, those results indicate that the information state marked by different linguistic cues (such as pitch accent, syntactic structure, and discourse context) can be immediately identified by the readers or listeners, and can interact with each other very quickly during on-line language comprehension.

The influence of IS on depth of language processing

In order to target specific aspects of language processing, some efforts have been devoted to investigating how IS influences the depth of language processing. In these studies, the depth of processing was indexed by whether people noticed anomalies and inconsistencies in sentences. The fact that people sometime fail to notice the anomalies and inconsistencies has been referred to as semantic illusions (Erickson and Mattson 1981). Interestingly, these studies showed that the occurrence of a semantic illusion became less frequent when the anomalous words were focused by an "it-cleft" structure (Bredart and Modolo 1988) or by a surface marking such as capitalization or underlining of the critical word (Bredart and Docquier 1989). In addition, Cutler and Fodor (1979) found that the reaction time to detect a phoneme target in the answer sentence was faster when the target phoneme occurred in focused words. Moreover, Sanford et al. (2009) found that the use of cleft sentences increases the processing efficiency of references. The difference in processing focused and non-focused information implies that IS modulates language processing by allocating more attentional resources to the focused than non-focused information, and hence the focused information obtains deeper processing.

Similar findings were observed in studies using a text-change detection paradigm (Sturt et al. 2004). In this paradigm, the subjects are presented with consecutive presentations of sentences, where some words are changed during the second presentation. The subjects are required to read the two presentations and decide whether any word has been changed on the second display. It has been found that the change detection rate increased for sentences where the changed words were focused by an "it-cleft" structure (Sturt et al. 2004), by a wh-word in the context (Ward and Sturt 2007), by italicization for the written words, and by accentuation for the spoken words (Sanford et al. 2006). The different detection rates indicate that not all information is processed equally. The changed words that are more likely to be detected must have been encoded in more detail and thus processed more deeply.

Overall, by comparing task performances between focused and non-focused information, these studies demonstrate that IS modulates the degree of elaboration of the language input. This modulation has been attributed to the assumption that focused information attracts more attentional resources and thus it is processed in more detail than non-focused information. However, all the above-mentioned studies employed explicit tasks (to judge the semantic appropriateness of sentences or to detect the changes of words in sentences), which might have involved different processes (e.g., decision making) from natural language comprehension. Given that no explicit task is needed in ERP technique, the influence of IS on the depth of language processing can be further examined using ERPs.

In two recent studies, we manipulated the semantic appropriateness of focused and non-focused words in question–answer pairs (e.g., e.g., *What vegetables did mum buy for dinner?/Who bought vegetables for dinner? Mum bought* ***eggplant/beef*** *for dinner.* The critical words are in boldface), and compared the ERP effects in response to the semantic inappropriateness between the focus and non-focus conditions. We found that the semantic inappropriateness (e.g., beef is not a kind of vegetables) of the focused words (i.e., in the *what-question* context) elicited a clear N400 effect, whereas a reduced N400 effect was elicited when the words were in non-focus position (i.e., in the *who-question* context) (Wang et al. 2009). Moreover, when the focused or non-focused words were realized with or without pitch accent in Dutch spoken language, the accented and focused words produced the largest N400 effect, with no difference in the N400 effect among the other conditions (Wang et al. 2011). In addition, Li and colleagues also found that the semantically inappropriate words elicited a N400 effect when the corresponding words were accented; however, no significant difference was observed between the inappropriate and appropriate words when they were de-accented (Li and Ren 2012). Taken together, these three studies directly demonstrate the influence of IS on the depth of semantic processing.

Another study further investigated whether IS also modulates the depth of syntactic processing (Wang et al. 2012). In this study, subtle (number agreement) or salient (phrase structure) syntactic violations were placed either in focus or out of focus in question–answer pairs. P600 effects to these violations were taken as reflections of the depth of syntactic processing. For subtle violations, a P600 effect was observed in the focus condition, but not in the non-focus condition. For salient violations, comparable P600 effects were found in both conditions. These results indicate that IS can modulate the depth of syntactic processing, but that this effect depends on the salience of the information: when subtle violations are not in focus, they are processed less elaborately, but the modulation of IS is overridden by the information saliency. The salience-dependency of IS modulation was also demonstrated in a study where the emotional salience and information status of words were manipulated in question–answer pairs (Wang et al. 2013). In this study, emotional and neutral information (e.g., *The principal made an* ***awful/excellent/general*** *evaluation.* The critical words are in boldface, which are negative, positive and neutral respectively) were marked to be focus (e.g., *What kind of evaluation did the principal make?*) or non-focus (e.g., *Who made an evaluation?*) in question–answer pairs. While non-focused neutral words elicited a larger N400 than focused neutral words, no N400 effect was found between the focused and non-focused emotional words over right posterior regions. The results elucidate the extent of influence of IS on language processing.

Although previous studies have clearly demonstrated the influence of IS on both semantic and syntactic analysis of language input, the link between IS and attention remains to be proven. Given its high spatial resolution, we used the fMRI technique to examine the underlying neural substrates involved in IS-based modulation of semantic processing (Kristensen et al. 2012). In this study, the brain regions that are sensitive to pitch accent and semantic congruence were identified. We observed that pitch accent activated a bilateral superior and inferior parietal context, which has been identified as a part of domain-general attention network. In addition, an interaction between pitch accent and semantic congruence was revealed in bilateral inferior parietal regions: a larger activation was found for accented than for unaccented words only when the words were incongruent, while equally large activation was observed between the accented and unaccented words in the congruent conditions. The results indicate that pitch accent, as a marker of IS, signals the saliency of focused words and thereby recruiting a domain-general attention network in the service of more extended processing of the most relevant information. In addition, semantic incongruence caused a greater LIFG activation only when the words were accented. Given that increased LIFG activation can be interpreted as increased processing/unification load, the lack of the LIFG activation for the incongruence words with no pitch accent indicates that they were less attended to and shallowly processed.

Overall, existing studies suggest that IS triggers attentional systems according to information saliency, in order to avoid partial or incomplete processing of language inputs. Meanwhile, the IS modulation of language processing also indicates that not all of language input is processed to the same extent. This might be the reason that people use IS to guide others' attention towards the most important information. Next, we will briefly discuss the interaction between attention and language processing in general.

Attention and language processing

The involvement of attention network indicates that language processing is not modular, as has been proposed previously (Fodor 1983). The modularity of language processing was best demonstrated by the relationship between syntactic and semantic operations, i.e., whether syntactic computations precede semantic processing. In classical models of sentence comprehension, the two-stage model (Frazier and Fodor 1978) has argued for separate syntactic and semantic parsing. It was further claimed that syntactic analysis is an "automatic" processing that does not need to engage attentional control, as supported by the early ERP effect named early left anterior negativity (Hahne and Friederici 1999; Friederici and Kotz 2003). It should be noted that an interactive view between syntactic and semantic processing has also been proposed (MacDonald 1993), and that the syntactic analysis has been considered to be a controlled process (King and Just 1991). In addition, the modularity of language processing was reinforced by the view that language is immune to other cognitive systems (Fodor 1983). This notion became unpopular since the demonstrations of language processing interacting with other modules such as music and vision (Jackendoff 2002).

On the other hand, most attention-related studies have focused on the perception of visual or auditory stimuli (Corbetta et al. 2000; Petersen and Posner 2012; Posner and Petersen 1990), which might be difficult to apply to complex systems like language. In the framework proposed by Petersen and Posner (2012) and Posner and Petersen (1990), three independent but related networks were defined in the attention system: alerting, orienting, and executive control. The alerting network is used to produce and maintain optimal vigilance and performance during tasks. The orienting network guides our focus toward selective and salient inputs (e.g., modality, location). The executive control network deals with conflict among different neural systems competing for control. On the basis of neuroimaging findings, two attention pathways: a top-down attention network and a bottom-up attention network have been identified (Corbetta et al. 2000). The top-down network mediates the allocation of top-down attention driven by knowledge, expectations or current goals. It involves a dorsal fronto-parietal network, including the intraparietal sulcus (IPS) and superior parietal lobe (SPL), as well as the dorsal frontal cortex along the precentral sulcus, near or including the frontal eye field (FEF). The bottom-up network mediates bottom-up attention driven by relevant stimuli, especially unexpected and novel ones. It involves ventral fronto-parietal network, including the inferior parietal lobe (IPL) and the ventral frontal cortex, including parts of the middle frontal gyrus (MFG), the inferior frontal gyrus (IFG) and the anterior insula. Besides these two networks, subcortical structures such as superior colliculus and pulvinar nucleus of the thalamus are also shown to be important in coordinating attention (Shipp 2004).

The two framworks proposed by Posner and Petersen (1990) and Corbetta et al. (2000) focused on different aspects of attention, i.e., cognitive and neurobiological architecture of attention. Regardless, these models were mostly built on the basis of findings of perception of different objects, specific characteristic of object, or spatial location. Although the role of attention in language compression has been recognized previously (e.g., Huettig et al. 2011; Bialystok 1999; Bialystok and Viswanathan 2009), little is known regarding the neural mechanisms involved in language comprehension under the influence of attention. Therefore, we call for more systematic studies on the interaction between attention and language. The described attention-related models set up a platform to test the attention component/network involved in language processing. In the end, we will bring up some possible directions for future studies.

Future directions

Existing studies on the role of IS in language processing have demonstrated that some linguistic markers can be used to modulate attention allocation. However, an open question is whether different IS makers recruit a general or marker-specific attention network. Our previous study has shown that pitch accent activated both top-down and bottom-up attention networks (Kristensen et al. 2012). IS can be invoked in various ways, such as question context, syntactic structure (e.g., "it-cleft" structure), focus particles (e.g., "only", "even"). While some types of focus marking (such as context and syntax) do not typically change the propositional content of the marked information, focus particles interface with semantic interpretations. Thus, it will be interesting to see whether other types of IS marking also recruit the same attention network. Also, it will be interesting to further specify which attention component or attention network is sensitive to a certain type of linguistic marker.

In addition to modulating attentional resources, IS-markers also convey semantic meaning. For instance, pitch accent was shown to modulate attention for highly expected new information whereas it mostly signified the information status of lowly expected new information (Li and Yang 2013). In addition, Chen, Li, and Yang (2012) found that the reading pattern of focus versus non-focus invoked by Chinese IS-marker "shi" (which conveys a similar meaning as "it-cleft" in English) was different from that of new versus given imposed by sentence context: while people spent longer time reading new information, they spent shorter time reading focused information. Moreover, the comparison of focus versus non-focus elicited a different ERP effect (i.e., a positive effect) from the comparison of new versus given (i.e., an N400 effect) (Chen et al. 2014). The results suggest that attention allocation and lexical retrieval were involved respectively in the two types of IS-marking. Likewise, some focus-marking particles (e.g., *only*, *even*) could convey semantic meaning (e.g., *exclusive*, *additive*) over and above attention modulation (Gotzner et al. 2013). As previous studies have shown that the role of IS interacts with some aspects of words (e.g., the emotional valence of words in the study by Wang et al. 2013), it is important to determine to what extent and under what circumstance does IS modulate attention.

Another interesting question is how non-linguistic devices affect attention allocation and further influence language comprehension. For instance, beat gesture is a rapid movement of the hand usually in an up and down manner, produced with the rhythm of the concurrent speech (McNeill 1992). It is often used to highlight new or contrastive information but conveys no semantic content (Alibali et al. 2001). Also, beat gesture and pitch accent were shown to be closely related in time and in function (Krahmer and Swerts 2007; Leonard and Cummins 2010). Our recent ERP study found that beat gesture and pitch accent independently modulate semantic processing in isolated sentences, implying their independent roles in capturing attention (Wang and Chu 2013). Except beat gesture, facial expressions (such as nodding and eyebrow raising) also affect speech intelligibility (Al Moubayed and Beskow 2009). Therefore, the interaction between linguistic and non-linguistic devices in modulating attention requires further investigation.

Overall, the studies on cognition-general and language-specific attention will allow us to study language processing within a broad cognitive framework and to better understand the nature of language processing.

Acknowledgments This research was supported the Natural Science Foundation of China (Grant Number: 31200849).

References

Al Moubayed S, Beskow J (2009) Effects of visual prominence cues on speech intelligibility. Paper presented at the Proceedings of the international conference on auditory visual speech processing AVSP

Alibali MW, Heath DC, Myers HJ (2001) Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. J Mem Lang 44(2):169–188. doi:10.1006/jmla.2000.2752

Bastiaansen MC, Van Der Linden M, Ter Keurs M, Dijkstra T, Hagoort P (2005) Theta responses are involved in lexical—semantic retrieval during language processing. J Cogn Neurosci 17(3):530–541

Bastiaansen M, Oostenveld R, Jensen O, Hagoort P (2008) I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. Brain Lang 106(1):15–28

Bialystok E (1999) Cognitive complexity and attentional control in the bilingual mind. Child Dev 70(3):636–644

Bialystok E, Viswanathan M (2009) Components of executive control with advantages for bilingual children in two cultures. Cognition 112(3):494–500. doi:10.1016/j.cognition.2009.06.014

Birch S, Clifton C (1995) Focus, accent, and argument structure: effects on language comprehension. Lang Speech 38(4):365–391. doi:10.1177/002383099503800403

Bock JK, Mazzella JR (1983) Intonational marking of given and new information: some consequences for comprehension. Mem Cogn 11(1):64–76

Bögels S, Schriefers H, Vonk W, Chwilla DJ (2011) Pitch accents in context: how listeners process accentuation in referential communication. Neuropsychologia 49(7):2022–2036. doi:10.1016/j.neuropsychologia.2011.03.032

Bredart S, Docquier M (1989) The Moses illusion: a follow-up on the focalization effect. Curr Psychol Cogn 9:357–362

Bredart S, Modolo K (1988) Moses strikes again: focalization effect on a semantic illusion. Acta Psychol 67(2):135–144

Chen L, Li X, Yang Y (2012) Focus, newness and their combination: processing of information structure in discourse. PLoS One 7(8):e42533

Chen L, Wang L, Yang Y (2014) Distinguish between focus and newness: an ERP study. J Neurolinguistics 31:28–41. doi:10.1016/j.jneuroling.2014.06.002

Chomsky N (1965) Aspects of the theory of syntax. MIT, Cambridge

Corbetta M, Kincade JM, Ollinger JM, McAvoy MP, Shulman GL (2000) Voluntary orienting is dissociated from target detection in human posterior parietal cortex. Nat Neurosci 3(3):292–297. doi:10.1038/73009

Corbetta M, Kincade JM, Shulman GL (2002) Neural systems for visual orienting and their relationships to spatial working

memory. J Cogn Neurosci 14(3):508–523. doi:10.1162/089892902317362029
Cowles HW (2012) The psychology of information structure. Expr Inf Struct 5:287
Cowles HW, Kluender R, Kutas M, Polinsky M (2007) Violations of information structure: an electrophysiological study of answers to wh-questions. Brain Lang 102(3):228–242
Cutler A, Fodor JA (1979) Semantic focus and sentence comprehension. Cognition 7(1):49–59
Dahan D, Tanenhaus MK, Chambers CG (2002) Accent and reference resolution in spoken-language comprehension. J Mem Lang 47(2):292–314
Dimitrova DV, Stowe LA, Redeker G, Hoeks JC (2012) Less is not more: neural responses to missing and superfluous accents in context. J Cogn Neurosci 24(12):2400–2418
Erickson TD, Mattson ME (1981) From words to meaning: a semantic illusion. J Verbal Learn Verbal Behav 20(5):540–551
Erteschik-Shir N (2007) Information structure: the syntax–discourse interface, vol 3. Oxford University Press, Oxford
Fodor J (1983) The modularity of mind. The MIT Press, Cambridge
Frazier L, Fodor JD (1978) The sausage machine: a new two-stage parsing model. Cognition 6(4):291–325
Friederici AD, Kotz SA (2003) The brain basis of syntactic processes: functional imaging and lesion studies. Neuroimage 20, Supplement 1(0) :S8–S17. doi:10.1016/j.neuroimage.2003.09.003
Gotzner N, Spalek K, Wartenburger I (2013) How pitch accents and focus particles affect the recognition of contextual alternatives. Paper presented at the Proceedings of the 35th Annual Conference of the Cognitive Science Society
Gűnther C, Maienborn C, Schopp A (1999) The processing of information structure. In: Bosch P, van der Sandt RA (eds) Focus: linguistic, cognitive, and computational perspectives. Cambridge University press, Cambridge, pp 18–43
Gussenhoven C (2008) Notions and subnotions in information structure. Acta Linguist Hung 55(3–4):381–395. doi:10.1556/ALing.55.2008.3-4.11
Hahne A, Friederici AD (1999) Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. J Cogn Neurosci 11(2):194–205. doi:10.1162/089892999563328
Halliday MAK (1967) Notes on transitivity and theme in English: part 2. J Linguist 3(02):199–244. doi:10.1017/S0022226700016613
Heim S, Alter K (2006) Prosodic pitch accents in language comprehension and production: ERP data and acoustic analyses. Acta Neurobiol Exp 66(1):55–68
Heim S, Alter K (2007) Focus on focus: the brain's electrophysiological responses to focus particles and accents in German. In: Späth A (ed) Language context and cognition: interfaces and interface conditions. De Gruyter, Berlin
Hruska C, Alter K (2004) Prosody in dialogues and single sentences: how prosody can influence speech perception. In: Steube A (ed) Information structure: theoretical and empirical aspects. Walter de Gruyter, Berlin, pp 221–226
Huettig F, Rommers J, Meyer AS (2011) Using the visual world paradigm to study language processing: a review and critical evaluation. Acta Psychol 137(2):151–171. doi:10.1016/j.actpsy.2010.11.003
Ito K, Garnsey SM (2004) Brain responses to focus-related prosodic mismatch in Japanese. Paper presented at the Speech Prosody 2004, International Conference
Jackendoff R (2002) Foundations of language: brain, meaning, grammar, evolution. Oxford University Press, New York
Jensen O, Bonnefond M, VanRullen R (2012) An oscillatory mechanism for prioritizing salient unattended stimuli. Trends Cogn Sci 16(4):200–206. doi:10.1016/j.tics.2012.03.002
Johnson SM, Breen M, Clifton CJR, Morris J (2003) ERP investigation of prosodic and semantic focus. Poster presented at Cognitive Neuroscience, New York City
Karttunen L, Stanley P (1979) Coventional implicature, vol 11. Academic Press, New York
King J, Just MA (1991) Individual differences in syntactic processing: the role of working memory. J Mem Lang 30(5):580–602
Krahmer E, Swerts M (2007) The effects of visual beats on prosodic prominence: acoustic analyses, auditory perception and visual perception. J Mem Lang 57(3):396–414. doi:10.1016/j.jml.2007.06.005
Kristensen LB, Wang L, Petersson KM, Hagoort P (2012) The interface between language and attention: prosodic focus marking recruits a general attention network in spoken language comprehension. Cereb Cortex. doi:10.1093/cercor/bhs164
Kutas M, Federmeier KD (2011) Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu Rev Psychol 62(1):621–647
Lambrecht K (1996) Information structure and sentence form: topic, focus, and the mental representations of discourse referents, vol 71. Cambridge University Press, Cambridge
Leonard T, Cummins F (2010) The temporal relation between beat gestures and speech. Lang Cogn Process 1–15. doi:10.1080/01690965.2010.500218
Li X, Ren G-Q (2012) How and when accentuation influences temporally selective attention and subsequent semantic processing during on-line spoken language comprehension: an ERP study. Neuropsychologia 50(8):1882–1894
Li X, Yang Y (2013) How long-term memory and accentuation interact during spoken language comprehension. Neuropsychologia 51(5):967–978
Li X, Hagoort P, Yang Y (2008) Event-related potential evidence on the influence of accentuation in spoken discourse comprehension in Chinese. J Cogn Neurosci 20(5):906–915. doi:10.1162/jocn.2008.20512
Li X, Chen Y, Yang Y (2011) Immediate integration of different types of prosodic information during on-line spoken language comprehension: an ERP study. Brain Res 1386:139–152. doi:10.1016/j.brainres.2011.02.051
MacDonald MC (1993) The interaction of lexical and syntactic ambiguity. J Mem Lang 32(5):692–715
Magne C, Astesano C, Lacheret-Dujour A, Morel M, Alter K, Besson M (2005) On-line processing of "pop-out" words in spoken French dialogues. J Cogn Neurosci 17(5):740–756. doi:10.1162/0898929053747667
McNeill D (1992) Hand and mind: what gestures reveal about thought. University of Chicago Press, Chicago
Paterson KB, Liversedge SP, Filik R, Juhasz BJ, White SJ, Rayner K (2007) Focus identification during sentence comprehension: evidence from eye movements. Q J Exp Psychol 60(10):1423–1445
Petersen S, Posner M (2012) The attention system of the human brain: 20 years after. Annu Rev Neurosci 35(1):73–89
Posner MI, Petersen SE (1990) The attention system of the human brain. Annu Rev Neurosci 13:25–42
Sanford AJS, Sanford AJ, Molle J, Emmott C (2006) Shallow processing and attention capture in written and spoken discourse. Discourse Process 42(2):109–130
Sanford AJ, Price J, Sanford AJ (2009) Enhancement and suppression effects resulting from information structuring in sentences. Mem Cogn 37(6):880–888
Shahin AJ, Picton TW, Miller LM (2009) Brain oscillations during semantic evaluation of speech. Brain Cogn 70(3):259–266
Shipp S (2004) The brain circuitry of attention. Trends Cogn Sci 8(5):223–230

Steedman M (1991) Structure and intonation. Language 67(2):260–296

Sturt P, Sanford A, Stewart A, Dawydiak E (2004) Linguistic focus and good-enough representations: an application of the change-detection paradigm. Psychon Bull Rev 11(5):882–888. doi:10.3758/bf03196716

Swerts M, Krahmer E, Avesani C (2002) Prosodic marking of information status in Dutch and Italian: a comparative analysis. J Phenetics 30(4):629–654

Terken J, Nooteboom SG (1987) Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. Lang Cogn Process 2(3):145–163

Toepel U, Pannekamp A, Alter K (2007) Catching the news: processing strategies in listening to dialogs as measured by ERPs. Behav Brain Funct 3(1):53

Wang L, Chu M (2013) The role of beat gesture and pitch accent in semantic processing: an ERP study. Neuropsychologia 51(13):2847–2855

Wang L, Hagoort P, Yang Y (2009) Semantic illusion depends on information structure: ERP evidence. Brain Res 1282:50–56

Wang L, Bastiaansen M, Yang Y, Hagoort P (2011) The influence of information structure on the depth of semantic processing: how focus and pitch accent determine the size of the N400 effect. Neuropsychologia 49(5):813–820

Wang L, Bastiaansen M, Yang Y, Hagoort P (2012) Information structure influences depth of syntactic processing: event-related potential evidence for the Chomsky illusion. PLoS One 7(10):e47917. doi:10.1371/journal.pone.0047917

Wang L, Bastiaansen M, Yang Y, Hagoort P (2013) ERP evidence on the interaction between information structure and emotional salience of words. Cogn Affect Behav Neurosci. doi:10.3758/s13415-012-0146-2

Ward P, Sturt P (2007) Linguistic focus and memory: an eye movement study. Mem Cogn 35(1):73–86