Title

WPP, No.111: Focus, prosody, and individual differences in “autistic” traits: Evidence from cross-modal semantic priming

Author
Bishop, Jason

Publication Date
2012-12-13

UCLA Working Papers in Phonetics, No.111 (2012), pp.1-26

Focus, prosody, and individual differences in "autistic" traits: Evidence from cross-modal semantic priming

Jason Bishop
(j.bishop@ucla.edu)

Abstract

The present study explored listeners' expectations about how prosodic prominence can be used to disambiguate information structure in English. In particular, the contribution of prenuclear accents to the prosodic disambiguation of the size of the focus constituent (broad VP vs. narrow object focus) in SVO constructions was tested using the cross-modal priming paradigm. In two experiments, listeners were presented with visual targets (e.g., *brunette*) following contrastively related primes (e.g., *blonde*), which were heard as objects in SVO sentences (e.g., *He kissed a blonde.*). In Experiment 1, listeners heard the sentences produced with a single pitch accent on the object, and the focus structure varied from broad VP focus to narrow object focus. No significant differences in priming patterns across conditions were found, supporting theories of Focus Projection (e.g., Selkirk 1995, Gussenhoven 1984), which predict prenuclear accents to be optional. In Experiment 2, the information structure of the sentences was held constant as narrow object focus, and their prosody varied with respect to the presence of a prenuclear pitch accent on the verb. For these narrow focus sentences, it was found that priming occurred only when the sentence lacked a prenuclear accent, suggesting that prenuclear pitch accents contribute meaningfully to the information structural contrast. Sensitivity to the prosodic manipulation, however, was found to be modulated by individual differences in listeners' "autistic" traits. The implications for on-line lexical processing and theories of the mapping between prosody and information structure are discussed.

1. Introduction

An important aspect of a sentence's meaning is the relation between its content and the larger discourse context assumed by a speaker and hearer. The primary linguistic mechanism mediating this relationship is known as information structure, a "packaging" of a sentence's information into categories such as "focus", "topic", and "given" (for recent summaries, see Féry and Krifka 2008, and Büring 2007). In many languages, it is clear that primary vehicle for information structure is prosody, although many details of the prosody-information structure relationship are far from understood. The present paper explores how a particular contrast along the dimension of focus is expressed prosodically in English. In particular, we examine English-speaking listeners' expectations regarding how prosodic prominence can distinguish the size of the focus constituent of a sentence, the consequences of these expectations for on-line processing, and individual differences in listeners' reliance on such expectations. Since it is central to several research programs in both linguistics and psycholinguistics, we begin by describing the information structural contrast of interest in some detail, and some important theoretical and experimental findings related to it. A novel experiment is then presented that probes listeners' expectations for how prenuclear accents contribute to the prosodic expression of the size of the focus constituent.

1.1 Focus and focus structure

The focus of a sentence contains the informative part of the sentence's content – that which cannot be inferred from the discourse. The inferable information, on the other hand, is regarded as "given". Thus, for any sentence, what is focused and what is given is dependent on the context in which it is uttered. For example, for the simple subject-verb-object (SVO) sentence in (2), what the focus will be is dependent on the discourse context in which it is uttered.

(1) a. What happened?
b. What did Robert do?
c. Who did Robert kiss?

(2) Robert kissed a blonde.

Although the basic semantic meaning in (2) does not vary across contexts (the lexical items, their referents, and the grammatical relations between them are the same), the information that is informative or highlighted differs. In the context of (1a), the sentence in (2) contains all new, informative information; in the context of (1b), which asks about an activity carried out by Robert, only the verb phrase (VP) "*kissed a blonde*" is informative; in the context of (1c), which asks for the object of a kissing event carried out by Robert, the informative part of the message is limited to the object "*the blonde*". In this sense, the focus of a sentence is fixed pragmatically by the discourse, and can be "broadly" on a large syntactic constituent like a sentence or VP, or "narrowly" on a single word or noun phrase (Ladd 1980). In terms of semantic interpretation, the influential theory of Alternative Semantics (Rooth 1992) claims that the presence of a focus feature on a constituent introduces a "focus semantic value" – understood as a set of alternative meanings for that constituent. Thus, in addition to the ordinary meaning of the sentence "*He kissed a blonde*", focus on the object "*blonde*" introduces additional propositional meanings such as "*He kissed a brunette*", "*He kissed a red-head*", "*He kissed Mary…*", and so on. In the case of focus that is broadly on the VP, the focus semantic value would include sentences where the VP is replaced with alternative VPs, such as "*He left*", "*He talked to Mary*", "*He made a phonecall*", etc.

While the distinction between broad and narrow focus represents a clear and categorical contrast in meaning, its prosodic realization has been a matter of much debate. In some linguistic models of the prosody-information structure interface, focus is conveyed through the location of the sentence's nuclear pitch accent – the last and usually most prominent pitch accent in a sentence. Thus, a production of the sentence in (2), with a nuclear pitch accent on "*Robert*" is said to be acceptable in all three contexts in (1). In general, such models recognize a class of prenuclear accents – accents preceding and usually less prominently perceived than the nuclear accent – but neither they, nor other aspects of sentence prosody encode focus structure. This is the position taken by Ladd (1996), who claimed that any differences speakers might be observed to produce for different focus structures are best regarded as "paralinguistic emphasis" – the result of (presumably language independent) use of phonetic prominence that is not part of linguistic knowledge. Büring (2007) takes a similar position, regarding prenuclear accents as "ornamental" – i.e., optional with respect to meaning, but possibly necessary for phonological purposes, such as the maintenance of rhythm.

The size of the focus constituent in SVO sentences is also expected to be prosodically ambiguous by an extremely influential class of model, namely theories of "focus projection", as represented by Selkirk (1984/1995) and Gussenhoven (1984/1999). In Gussenhoven's model, nuclear accents mark semantic constituents, such as subjects, predicates and their arguments. A nuclear accent marking an internal argument of a verb is able to "project" focus up to the larger predicate, thus allowing a nuclear accent on an object to be acceptable for either broad VP or narrow object focus. Prenuclear accents, on the other hand, are added optionally, and subsequent to the nuclear accent, by phonological rule. In the focus projection model proposed by Selkirk, accenting marks syntactic rather than semantic constituents, and focus is allowed to percolate up the syntactic structure to mark larger syntactic phrases via a set of rules, resulting in the same ambiguity when the nuclear accent falls on the object. One way Selkirk's model differs from other models, however, is that prenuclear accents are not entirely unpredictable. Because a separate stipulation in her theory holds that accents not marking the primary focus are interpreted as marking information as "new", it predicts that narrow object focus should contain no prenuclear accent on the verb. Thus for Selkirk, prenuclear accents do have a certain semantic significance, and are predicted to be completely optional only under broad focus, a frequently overlooked aspect of her theory (but see Welby 2003). However, suffice it to say that theoretical work, emphasizes the fact the nuclear accent is the most reliable correlate of focus location, and when it does not vary, a high degree of ambiguity is expected to result. However, as is evidenced below, experimental phonetic and perception evidence has suggested less ambiguity in the minds of speakers and listeners.

1.2 Prosodic Realization of broad and narrow focus

Subsequent to most of the theoretical work described above, a number of phonetic studies have accumulated in English and closely related German and Dutch that suggest speakers can disambiguate their productions of broad and narrow focus. One of the first indications of this was reported by Gussenhoven (1983), who acquired productions of broad VP and narrow object focus sentences and presented them to listeners out of context. Although Gussenhoven did not provide an acoustical analysis of these production data, differences are inferable from the fact that listeners reported hearing them. These differences had to do with the prominence of the verbs, such that verbs in SVO sentences that had been produced in VP focus contexts were rated as sounding more prominent by listeners.

Later work has confirmed this initial observation, but shown that the pattern is subject to much variation. Phonetic production studies have often concentrated on f0 as a phonetic measure, and reported that compared with the same object within a broader VP or sentence focus, an object under narrow focus will be produced with higher f0. Importantly, however, this boosting of the pitch peak may be accomplished by the speaker in one of two basic ways: either directly, by way of increasing f0 of on the object (Eady and Cooper 1986, Eady et al. 1986), or more indirectly, by additionally decreasing prominence on prenuclear material (Xu and Xu 2005; see also Baumann et al. 2006 and Baumann et al. 2008 for German). This seems to suggest that speakers encode the contrast prosodically by means of relative prominence manipulation, a fact further supported by the finding that speakers will sometimes suppress post-nuclear prominence as well (Xu and Xu 2005 and Hannsen et al. 2008 for Dutch). Thus phonetic prominence relations seems to convey the contrast, at least in the minds of speakers, and details of the of this relative prominence relation may be subject to variation in implementation strategies.

Careful inspection of other phonetic parameters of prominence has further confirmed this use of relative prominence to convey the size of a focus constituent. It has been shown, for example, that speakers can encode the distinction using segmental or word durations (Sityaev and House 2003, Baumann et al. 2006, Hannsen et al. 2008), and this, like f0, may be implemented indirectly by some speakers (Eady and Cooper 1986, Eady et al. 1986, Breen et al. 2010). In fact, in a recent study, Breen et al. (2010) demonstrate that previous studies, which have generally employed small numbers of speakers and fairly uncommunicative speech situations, may have considerably overestimated the ambiguity of the focus size contrast. In their large scale production study, the authors found that a statistical model could correctly classify native English-speaker productions as broad (sentence) or narrow object focus with a high level of success, somewhat more so in the case of narrow focus. To achieve this, however, models required not only multiple prominence-lending cues – f0, duration, and (in fact, most importantly), intensity – but they needed to know them for prenuclear words in the sentence. Thus, accurate classification depended crucially on multiple relative measures of prominence.

1.3 Listeners' use of prosodic cues to focus size

Rather surprisingly, some perception studies investigating the prosodic cues just described have failed to find evidence that listeners use them. Gussenhoven (1983), for example, although he found VP focus productions to have more prominent verbs, found listeners to ignore such prominence when matching SVO sentences with broad or narrow focus contexts. Using an appropriateness rating task, this same ambivalence was replicated by Welby (2003), who additionally found listeners to lack any preference for a particular type of nuclear accent on the object; listeners rated as equally acceptable sentences with a phonetically prominent ToBI[1] L+H* accent or the more neutral H* accent for broad VP focus (she did not test this distinction for narrow object focus). Finally, Birch and Clifton (1995) also tested listeners for a preference for the presence versus absence of a prenuclear accent, though only for broad VP focus, and their results were more mixed. While they found prenuclear accents irrelevant to listeners' ratings when the task was to judge whether the sentence "made sense" in the context of the question, they sound small but significant differences when the task was to judge the intonational felicity specifically. These differences were in the direction of preferring the broad VP focus sentences in their study to bear a prenuclear accent on the verb.

The results of Birch and Clifton's study suggest the possibility that drawing attention to the prosodic ambiguity may be important for eliciting listeners' knowledge about how the contrast can be distinguished. This effect has been shown for speakers (e.g., Snedecker and Trueswell 2003, Jun 2010), a fact which is relevant for the stimuli used in the perception studies just reviewed. Further evidence for this comes from Breen et al. (2010), whose study, reviewed above, reports results of perception by human listeners in addition to their disciminant anaysis. They found that listeners' successful classification generally mirrored that of the acoustic cues in a task that made them aware of the prosodic ambiguity. This indicates that when speakers were trying to disambiguate, and listeners were aware of the ambiguity, successful transmission of the meaning contrast took place via prosody. Interestingly, however, they found that not all listeners in their study were able to do this above chance level, indicating individual differences in the ability to attend to and correctly interpret the cues even when they were present.

[1] Unless otherwise stated, the intonational labels used in this paper refer to Tones and Break Indices conventions for Mainstream American English (Beckman and Hirschberg 1994).

Finally, tasks that do not rely on metalinguistic appropriateness ratings or context matching seem to produce more reliable differences in listeners' behavior. For example, using a task similar to the psychophysical method-of-adjustment, Rump and Collier (1996) asked Dutch listeners to adjust the height of prenuclear and nuclear accent peaks for synthetic versions of sentences. Listeners were to make these sentences appropriate for different focus structures, and it was found that their adjustments varied systematically, such that narrow object foci required lower prenuclear accents and higher nuclear accents, and broad focus required the opposite pattern. These results, in line with production patterns, suggest that listeners have clear expectations about "prototypical" prosodic realizations of broad and narrow focus. Such expectations were even found to have top-down effects on prominence perception by Bishop (2012). In that study, native English-speaking listeners gave subjective prominence ratings for verbs and objects in SVO sentences presented in either broad or narrow focus contexts. Although the same production of each sentence was always heard, and only the context varied, listeners reported hearing the objects as more prominent (and the verbs as less prominent) in the narrow focus context. Thus, in both of these experiments, which drew attention to the prosody of sentences (and probably the ambiguity as well), listeners showed clear preferences that were in line with patterns reported for speakers' productions.

To summarize, the broad versus narrow focus contrast in SVO constructions in English and closely related languages is predicted by several prominent theories to be prosodically ambiguous. However, a considerable amount of phonetic evidence suggests that narrowly focused objects are more phonetically prominent – relative to prenuclear material – than the same object situated within a larger focus constituent. Whether or not listeners use the phonetic patterns in the expected way has been less clear, but may depend on whether the speaker has encoded them sufficiently in production (a property of the stimulus), whether the listeners' attention is drawn to the prosodic ambiguity (a property of the task), and there might also be individual differences involving the attention to, and interpretation of, prosodic cues even when they are present (a property of the listener). We now turn to the present study, which used cross-modal associative priming to examine whether on-line processing mechanisms might reveal listeners' prosodic expectations more clearly than some off-line tasks have been able to do.

1.4 Cross-Modal Associative Priming

In the cross-modal associative priming paradigm, a listener is auditorily presented with a word, either in isolation or embedded in a sentence, and must then make a lexical decision about a related target word, subsequently presented visually. Priming occurs when the target word is recognized more quickly following a related prime word than following an unrelated control prime word. The use of cross-modal priming in the present study is motivated by recent work demonstrating that priming effects are highly dependent on factors related to sentence-level interpretation. In particular, Norris et al. (2006) show convincingly that semantic priming fails just in those cases where the associative relationship between prime and target is not supported by a larger sentence meaning (see also Tabossi 1996). This was also an emphasis in earlier work by Foss and Ross (1983), who argued that an "effective context" was necessary to activate sentence-level propositional meaning, which was in turn what actually led to priming.

Most pertinent to the present study is the finding that listeners use prosody to construct this effective context (Norris et al. 2006, Braun and Tagliapietra 2010). For example, Braun and

Tagliapietra (2010) presented Dutch-speaking listeners with primes such as *"flamingo"*, which occurred as objects in sentences such as *"In Florida he photographed a flamingo"*. These sentences were presented to listeners with either "neutral" prosody (a prenuclear accent on the subject, followed by a downstepped nuclear accent on the object) or with "contrastive" prosody (where the object bore a highly prominent H*L ToDI (Gussenhoven 2005) accent). They found that for contrastively related targets such as "*pelican*", priming only occurred when the prime was contrastively accented; for non-contrastively related targets (e.g., "*pink*"), however, priming occurred regardless of prosody, and was weaker.

Braun and Tagliapietra's findings are important because they show both that there is psychological reality to the notion of focus alternatives (i.e., to be contrastive is to activate alternatives), and that a particular prosody signals that focus meaning in the absence of explicit context. In the present study, our goal was to similarly probe listeners for such knowledge, but to do so by providing them with both the prosody and the context explicitly. This allows us to pair sentences with an unambiguous focus structure with different, putatively ambiguous, prosodic structures, and observe listeners' processing. To the extent that it is dependent on an appropriate match between semantic meaning and phonetic realization (Swinney 1979, Blutner and Sommer 1988) we expect priming effects to occur only when listeners' expectations are met.

A set of predictions about those expectations, derived from the results of production studies reviewed, is summarized in Table 1. The present investigation limited itself to testing the predictions shadowed in the table, which are the more contentious with respect to the theoretical literature. Experiment 1 tested whether a single pitch accent on the object in SVO sentences is equally acceptable to listeners in the context of a broad VP focus and a narrow object focus. Here, focus projection theories predict genuine ambiguity, although, production studies suggest that relative prominence is such that prenuclear material should be more prominent (i.e., accented), and thus we expect priming to occur only when there is a prenuclear accent on the verb. Experiment 2 tested whether prenuclear accents were truly optional for narrow object focus. This matter is a bit more contentious in the theoretical literature; while most accounts predict ambivalence in this case, Selkirk's (1995) theory predicts prenuclear accents to be infelicitous, since such accents are interpreted as marking new information. The production evidence suggests that relative prominence on an object should be high under narrow focus, and so also predicts prenuclear accentuation to be less felicitous. Before describing the two experiments that tested these hypotheses, however, there is one final topic to introduce, and that is the matter of listener-based variation.

Table 1. Predictions of priming patterns, based on listeners' preferences in production studies. # Indicates a contextually inappropriate prosodic realization.

	+ Prenuclear Accent	– Prenuclear Accent
VP focus	√	#
Object Focus	#	√

1.5 Individual Differences

Compared with the speaker, the listener as a source of variation has received much less attention in phonetic and psycholinguistic literature, although research is moving towards addressing this gap (e.g., Surprenant and Watson 2001; Makashay 2003, Stewart and Ota 2008, Cole et al. 2010, Yu 2010, Kong and Edwards 2011). Especially since listeners are known to vary in their ability to decode prosodic cues successfully (Breen et al. 2010), the present investigation attempted to explore sources of individual differences by including measures of "autistic"-like personality traits, motivated below.

Autistic traits are those behaviors and patterns of information processing associated with a clinical diagnosis with an Autism Spectrum Disorder. However, such traits – for example, non-holistic attentional focus, lack of social engagement, and poor communication skills – are known to occur to varying degrees in the neurotypical population as well. These traits are measured in non-clinical populations using the Autism Spectrum Quotient (AQ; Baron-Cohen et al. 2001), a non-diagnostic, self-administered questionnaire that divides autistic traits into five separate dimensions pertaining to *social skills*, *attention to detail*, *attention switching abilities*, *communication skills*, and imagination. Studies have shown the instrument, which is scored such that higher scores indicate more autistic traits, to have a high level of cross-cultural validity (Wakabayashi et al. 2006; Hoekstra et al 2008; Ruta et al. 2011; Sonié et al. 2012). It is also known that males generally score higher than females, scientists and mathematicians score higher than humanists (Baron-Cohen et al. 2001), and musicians with absolute pitch discrimination score higher than those without (Dohn, Garza-Villarreal, Heaton, Vuust 2012).

Important for our purposes, autistic traits have been shown relevant to predicting results in speech and language processing experiments. For example, Ota and Stewart (2008) found that high total AQ scores (i.e., sum of all five subscales) were associated with less perceptual shifting of segment identifications in the direction of real words compared with nonce words (i.e., the "Ganong effect", Ganong 1980), suggesting that autistic traits are associated with less integration of top-down lexical information. This result is echoed by Yu et al. (2011), who found high total AQ, to be associated with attenuated effects of phonotactics on segmental perception.

In a study examining perceptual compensation for coarticulatory effects (e.g., Mann 1980; Mann and Repp 1981; Mitterer 2006, Fowler 2006), Yu (2010) tested the influence of two factors on the perception of a fricative acoustically ambiguous between /s/ and /ʃ/. The first was an adjacent vowel (/ɑ/ versus /u/), and the second was the sex of the speaker. Yu found compensation for vowel context to be sensitive to autistic traits, in somewhat complicated ways. First, there was a significant interaction between the sex of the listener and overall AQ scores, such that, for women, higher total AQ was associated with greater compensation for vowel context compared with other subjects. However, in a subsequent analysis that included individual AQ subscales, it was found that higher scores on the communication subscale (i.e., worse communication skills) corresponded to less compensation for both vowel context and, to some extent, speaker sex (with no interaction with the sex of the listener). Thus, in addition lexical information, there is also evidence that the use of more local, syntagmatic information in speech is modulated by autistic traits, particularly along the dimension of communication skills.

Finally, it is known that the communication subscale of the AQ is a predictor of sensitivity to pragmatic "violations". In an ERP investigation, Nieuwland et al. (2010) asked subjects to read sentences that were either informative (*Some people have pets, which require good care.*) or uninformative (*Some people have lungs, which require good care.*), depending on whether the target word (underlined in these examples) should be trivially true via pragmatic implicature. Readers with low to mid-range communication scores (i.e., very good to average communication skills) exhibited the expected adverse brain response (i.e., the N400) following the target word in the uninformative sentences, while those with high communication scores showed no such effect. Interestingly, neither group showed an effect when the target word was not placed before a comma, (e.g., *Some people have <u>lungs</u> that require good care.*), suggesting that the violation may have depended on the target words being interpreted as phrase-final, and thus focused.

In summary, the AQ, particularly the communication subscale of the AQ, measures individual differences among listeners that have consequences for speech perception and linguistic processing. In particular, autistic traits seem to predict the extent to which listeners integrate incoming information with context, whether context is construed globally (the lexicon, semantic and pragmatic knowledge) or more locally (phonetic or phonological context). Because these are of prime importance to both the relation of information structure and prosody (and also cross-modal priming), autistic traits were included in the design of the two experiments below. Additionally, a measure of working memory capacity was also included, as this was found to be relevant in one of the studies reviewed (Yu et al. 2011), and is known from a large literature to be relevant to sentence processing (for recent summaries, see Fedorenko, Gibson, and Rohde 2006 and Swets, Desmet, Hambrick, and Ferreira 2007).

2. Experiments

2.1 Methods

2.1.1 Materials

The basic design of primes, targets and sentences for the two cross modal priming experiments was similar to that used in Braun and Tagliapietra (2010). The materials consisted of target words (e.g., *"pepper"*), and primes that were either related to the target or were unrelated control primes (e.g., *"salt"* and *"tape"*, respectively). Because it has been shown that contrastive associative relationships are most likely to facilitate priming in contrastive contexts (Braun and Tagliapietra 2010), all related primes were contrastive with the targets. The prime-target pairs were selected as follows. First, 32 English nouns, mostly monosyllables or disyllabic words with a strong-weak stress pattern were chosen to serve as the primes. These primes were then used in a web experiment to elicit contrastively related associates from 80 native English-speakers. These participants were presented with the 32 primes in frames such as "He didn't say "X", he said ___", to which they responded with the first word that came to mind. For each of the primes, the most frequent response was selected and used as the target for that prime (the mean association rate was 47.7% of responses; range 26.2% – 87.7%). Thirty-two simple SVO sentences were then constructed, in which the primes were to serve as the sentence-final objects. Care was taken so that, for each sentence, the object prime was the only word semantically related to the target. Thus 64 SVO sentences (one version of all 32 sentences containing the related prime, a second version containing the unrelated control prime), were produced by a male speaker of American English trained in phonetics and intonational phonology.

In order to create the two prosodic conditions, the 64 sentences were produced with two different intonational contours. First, a production was recorded in which the verb bore a prenuclear H* pitch accent with a following nuclear accent that was intended to be ambiguous between a H* and a !H*. This production was used as the prenuclear accented condition (see

Figure 1A). A second version of each of the sentences was then read without prenuclear accents on the verb (or on the subject, in most cases, a pronoun); the object was produced with a prominent L+H* nuclear accent (see Figure 1B).

In order to hold the acoustic information for the primes (i.e., the objects) themselves constant, while manipulating only prenuclear accentuation, the same recording of the object prime word was used in both prosodic conditions. This was accomplished by excising the production of the prime from the prenuclear accented condition and splicing it into the unaccented condition, replacing the original production of the L+H* object. Thus, the sentences in the prenuclear unaccented condition were original, unedited productions, the sentences in the prenuclear accented condition were manipulated versions. These manipulated version were now also of phonological structure L+H*, although the accent was generally phonetically less prominent than the one it replaced (Figure 1C). All experimental sentences (i.e., those containing related primes, and those containing unrelated control primes) were created in this way and saved as wav files.

To demonstrate that the prosodic manipulation resulted the intended accentuation contrast, acoustic measurements were carried out for verbs and objects in the final stimuli. Measurement criteria followed the recommendations in Turk, Nakai and Sugahara (2006). The mean values across items for the most common acoustic correlates of phonetic prominence are shown in Table 2. As can be seen, f0 values for verbs in the +prenuclear accent condition were significantly higher than those of the objects for both the test and control conditions (on average 110% the height of the object for each type of primes, consistent with a “falling hat pattern”, or H* H* sequence). In addition, verbs were considerably longer and of higher intensity when they were produced with a prenuclear accent. Thus, while the primes themselves were the same productions with the same absolute acoustic properties, they were of considerably different relative prominence across the prosodic conditions.

Finally, in order to create the information structural conditions in which these sentences would occur, lead-in questions were created. These were WH-questions such as *“What did Robert do?”* (in the case of VP focus) and *“What did Robert borrow?”* (in the case of object focus). In order to produce maximally contrastive contexts for the focused constituents, yes/no questions were additionally made to follow the WH questions. For example, the full question contexts were of the form *“What did Robert do after the party?...Did he leave?”* for VP focus, or *“Who did Robert kiss after the party?...Did he kiss Mary?”* for object focus. Question contexts, like the test sentences, were constrained such that only words unrelated to the targets could be used. All question contexts were produced and recorded by a female speaker of American English and were made to precede the SVO sentences in the stimuli, so that the SVO test sentences appeared as corrective answers to them. The full list of all materials used (sentences, test primes, control primes, and targets) is shown in the Appendix. In addition to these materials, there were also 96 filler sentences with filler primes and filler targets. 64 of the filler trials contained non-word targets; of the remaining 32 filler trials, half contained primes that were semantically unrelated to the target words, and half were related. In other respects, filler trials were the same as the experimental trials, with the same two prosodic versions of each. (For fillers, the prosodic conditions were both natural productions, with no splicing being done). An additional set of 6 filler sentences was also created to be used as items in a brief practice session to familiarize participants with the experimental task.

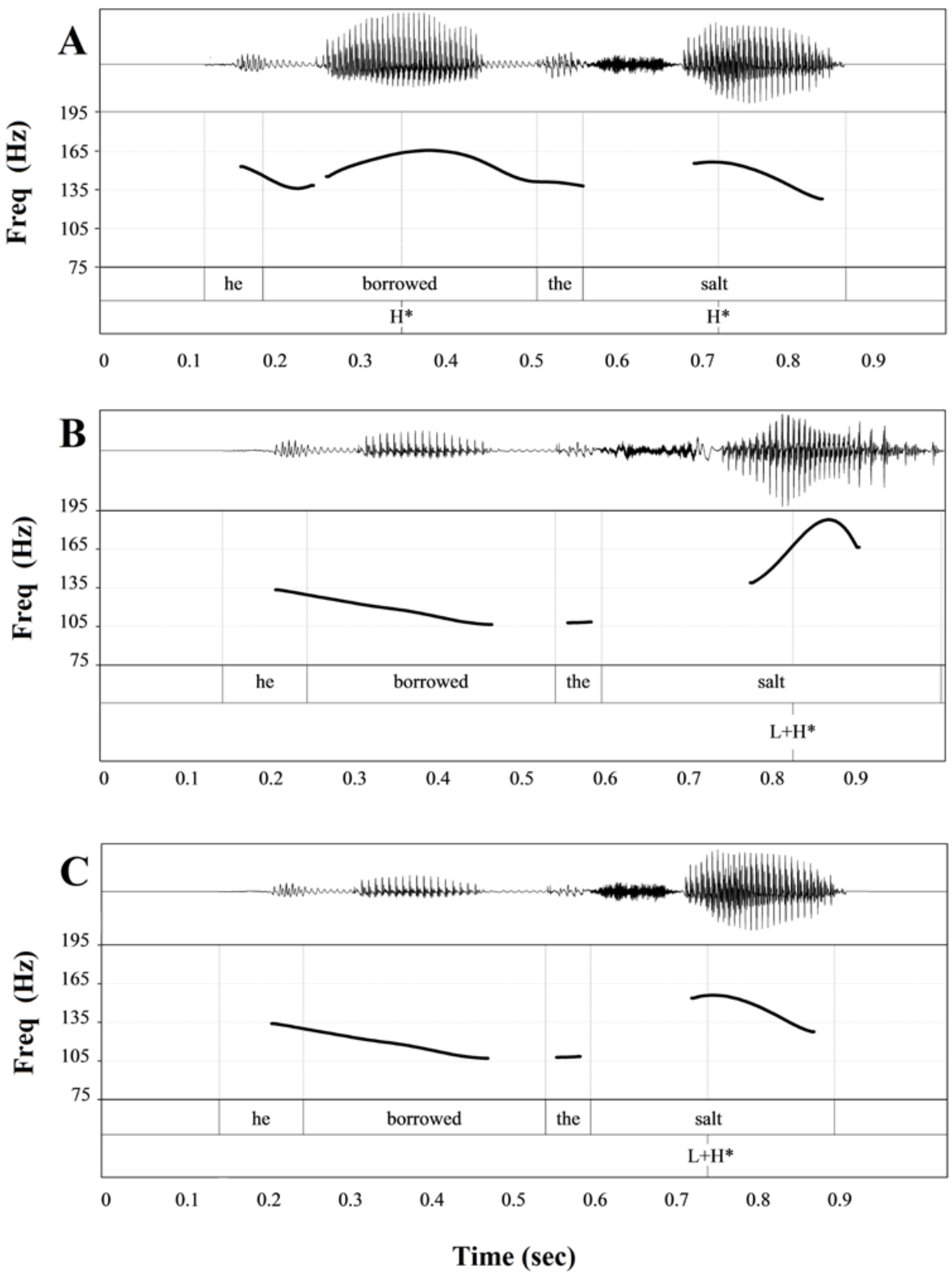

Figure 1. Example of an SVO test sentence in the two prosodic conditions. (A) shows the sentence produced with a prenuclear H* on the verb, used in the +prenuclear accent condition; (C) shows the sentence used for the –prenuclear accent condition, which was created by splicing the object "*salt*" from (A) into the production of the same sentence in (B).

Table 2. Acoustic properties of the verb and object in the two prosodic conditions for test and control sentences. The same object was used in both conditions, and so only one object is shown for each item type. Values shown are means with standard deviations in parentheses.

Test Items	+ Accent Verb	– Accent Verb	Object
Dur (ms)	294 (87)	257 (88)	466 (99)
Intensity (dB)	76.1 (2.9)	66.2 (3.3)	73.5 (3.1)
f0 min (Hz)	138 (12)	112 (7)	106 (9)
f0 max (Hz)	175 (19)	126 (10)	158 (13)
f0 range (Hz)	35 (14)	15 (8)	52 (14)

Control Items	+ Accent Verb	– Accent Verb	Object
Dur (ms)	304 (92)	260 (86)	520 (96)
Intensity (dB)	75.5 (3.4)	67 (4.9)	73.9 (3.6)
f0 min (Hz)	146 (13)	112 (9)	108 (12)
f0 max (Hz)	177 (16)	132 (14)	160 (12)
f0 range (Hz)	31 (15)	19 (9)	52 (15)

2.1.2 Participants

Ninety-two native English speakers, most of them members of the university community, participated as listeners in the lexical decision task in Experiment 1; Eighty-eight (different) native English speakers took part in Experiment 2 approximately 6 months later. None had participated in the web-based association experiment used for stimulus design, and none reported any history of a hearing, speech or communication disorder.

2.1.3 Procedure

Participants took part in a cross-modal lexical decision task, individually in a sound-attenuated booth. The auditory stimuli were played binaurally over Sony MDM headphones at a comfortable listening volume (held constant across participants). Visual targets appeared on a computer screen directly in front of the participant (in 72pt white font on a black background), immediately at the offset of the sentence-final primes. Participants were to push a ‘yes’ or ‘no’ key as quickly as possible to indicate that they recognized the word on the screen (the ‘yes’ key corresponded to the dominant hand for each participant). For Experiment 1, 4 lists were formed by taking all 32 test sentences and corresponding visual targets and rotating them through the two prime type conditions (related and unrelated prime) and the two information structural conditions (broad VP and narrow object focus contexts); all sentences contained prenuclearly accented verbs. For Experiment 2, 4 lists were created by rotating the same 32 sentences/visual targets through the two prime conditions and the two prosodic conditions (with or without a prenuclear accent on the verb); all sentences were presented in the narrow object focus contexts. Thus, in each experiment, there were 8 items per condition in each list; participants in each experiment were assigned (randomly) to one of the lists. A MATLAB script presented the stimuli in random order (different for each participant), and recorded responses and reaction times (RT). Following the lexical decision task, participants also completed (a) the Reading Span

Task (Daneman and Carpenter 1980; Unsworth et al. 2005) as a measure of verbal working memory capacity, and (b) the Autism Spectrum Quotient (Baron-Cohen et al. 2001). Participation in the entire experiment took approximately 40 minutes.

2.2 Results

2.2.1 Experiment 1

Reponses to experimental targets were considered errors if the participant failed to hit the 'yes' key, or their response was slower than 1800 ms. For the 92 subjects, this resulted in 63 errors (approximately 2.1% experimental trials), which were evenly distributed across the conditions and of no further interest. RTs for all correct responses falling within 2 standard deviations of the mean RT were analyzed (a total of 2819 observations) using mixed-effects linear regression using the *lmer* function in the *lme4* package (Bates and Maechler 2009) for R Statistics (R Development Core Team 2012).

The predictors that were of primary interest were (a) the linguistic predictor, the size of the focus constituent (VP versus object), (b) the prime type (related prime versus unrelated control prime), and, particularly, (c) their interaction. Also included were a number of stimulus-level variables known to be relevant to the lexical decision task: the CELEX log frequency of the target word, the length of the target word (in characters), and the reaction time to the preceding trial. Finally, the participant-level predictors were RSPAN score, score on the communication subscale of the AQ (henceforth AQ-Comm), sex, and the interaction of these predictors with each of the primary predictors (a-c). The initial model included all predictors as fixed-effect terms, as well as random intercepts for participant and item, and a by-participant random slope for trial. From this initial model, terms with a large p-value ($p > .1$) were then removed if it did not result in a significant decrease in model fit as assessed by a log-likelihood ratio test using the *anova* function in *R*. After removing non-contributing predictors in this way, the simplest model was refitted.

The resulting model is shown in Table 3. As expected given previous lexical decision studies, several stimulus-based predictors had a significant effect on RTs. In particular, RTs to targets were slower when the RT in the preceding trial was slower, when the target was longer or of low lexical frequency, and for trials that occurred earlier in the experiment. There were also main effects for (a) prime type, such that RTs were faster for targets following related primes, and (b) for focus, such that RTs were faster for targets that were interpreted as narrow object foci. Conspicuously absent from the model (because it was found to contribute nothing to model fit), was an interaction between prime type and focus, indicating that the priming effect observed was statistically equal for broad VP and narrow object focus.

The overall mean RTs for each of the experimental conditions is shown in Figure 2. Although there was no significant interaction between prime type and focus size, a tendency for priming of targets to be more effective when the target was narrowly focused was apparent (an average priming difference of 11.3 ms for object focus compared with 7.8 ms for VP focus). Nonetheless, the overall finding that priming occurred regardless of the size of the focus constituent indicates that sentences with a prenuclear accent on the verb were acceptable in both conditions.

Table 3. Results for fixed-effects factors for the model of reaction times in Experiment 1. Default level for the binary categorical factors is shown in italics.

	β	*SE (β)*	*t-value*	*p-value*
(Intercept)	523.740	15.334	34.15	< **.0001**
Trial	-0.298	0.058	-5.09	< **.0001**
Previous RT	0.045	0.008	5.41	< **.0001**
LogFreq of Target	-20.058	3.848	-5.21	< **.0001**
Target Length	3.664	1.795	2.04	< **.041**
Focus (*Narrow*)	-6.844	3.000	-2.28	< **.022**
Prime type (*Related*)	-7.744	2.996	-2.58	< **.009**

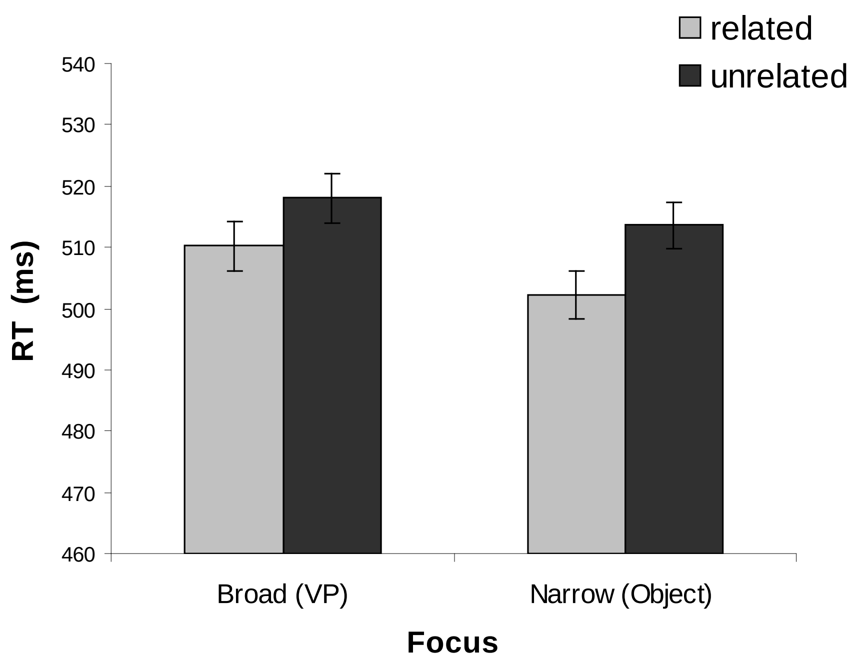

Figure 2. Mean reaction times for each of the information structural (focus) and prosodic conditions in Experiment 1. Error bars show standard error.

2.2.2 Experiment 2

Data from one participant were excluded due to very slow overall RT (fewer than 20% of responses were below 1800 ms). For the remaining subjects, error rate was determined as in Experiment 1, and was similar (2.5%, resulting in 2601 usable observations for the model). The modeling procedure and predictors were as in Experiment 1, with the exception that the linguistic predictor of primary interest was not the information structural status of the prime, but the sentence's prosodic structure (+/– prenuclear accent on the verb). This variable was included in the same two and three-way interactions as focus was in Experiment 1.

Table 4. Results for fixed-effects factors for the model of reaction times of participants' reaction times in the lexical decision experiment in Experiment 2. Default factor for the binary categorical factors is shown in italics.

	β	*SE (β)*	*t-value*	*p-value*
(Intercept)	741.809	42.690	17.38	< **.0001**
Trial	-0.324	0.010	-3.25	**.001**
Prev RT	0.014	0.006	2.22	**.027**
LogFreq of Target	-28.448	7.243	-3.93	< **.0001**
Target Length	4.281	3.385	1.27	.206
RSPAN	-2.489	0.723	-3.44	< **.001**
Prime Type(*Related*)	17.093	9.327	1.83	*.067*
Prominence(*High*)	10.796	9.307	1.16	.246
AQ-Comm	2.697	7.326	0.37	.713
Prime Type(*Related*)*Prominence(*High*)	-34.474	13.145	-2.62	**.009**
Prime Type(*Related*)*AQ-Comm	-6.876	3.444	-1.99	**.046**
Prominence(*High*) *AQ-Comm	-5.701	3.419	-1.67	*.096*
Prime Type(*Related*)*Prominence(*High*)*AQ-Comm	10.477	4.879	2.15	**.039**

The resulting model is shown in Table 4. The stimulus-based predictors (i.e., previous reaction time, log lexical frequency of the target, target length, and trial) all had the same effect on reaction times as in Experiment 1, significantly, with the exception of target length. Additionally, RSPAN scores were a robustly significant predictor, such that higher scores (reflecting greater working memory capacity) were associated with shorter RTs. There were also non-significant trends for both prime type and the prosodic manipulation, although both were in the opposite direction predicted; RTs were faster following primes that were either unrelated to the target, or were presented with low prosodic prominence.

The effects of these factors, however, are better understood in terms of their interactions with each other, and with participants' AQ-Comm scores. First and most important for our purposes, there was a significant two-way interaction between prime type and prosodic prominence, such that the facilitation of targets by primes occurred primarily when the sentence lacked a prenuclear accent on the verb. However, this pattern is further qualified by the significant three-way interaction between prime type, prosody, and AQ-Comm. The consequence of this interaction can be seen in Figure 3, which divides participants into two groups based on their AQ-Comm score. For participants that fall into the lower end of the spectrum of AQ-Comm scores (indicating good communication skills), we see the expected pattern: there is significant facilitation of targets following related primes relative to unrelated primes when the sentence lacked any prenuclear accent, and no effect (even a trend towards inhibition) when the sentence contained a prenuclearly accented verb. However, for those subjects on the higher end of the AQ-Comm spectrum (indicating poorer, more autistic-like communication skills), there are no clear differences between conditions, and indeed there is a trend in just the opposite direction: slight facilitation in the when sentences contained a prenuclear accent and slight inhibition when they did not. This inverse pattern is even more apparent in Figure 4, which shows priming effects binned by listeners' actual AQ-Comm scores (which ranged from 0 to 6 for these subjects, out of 10 possible points). Priming effects in the figure are determined by taking the average difference in RTs between targets following related and unrelated primes for each prosodic condition. As

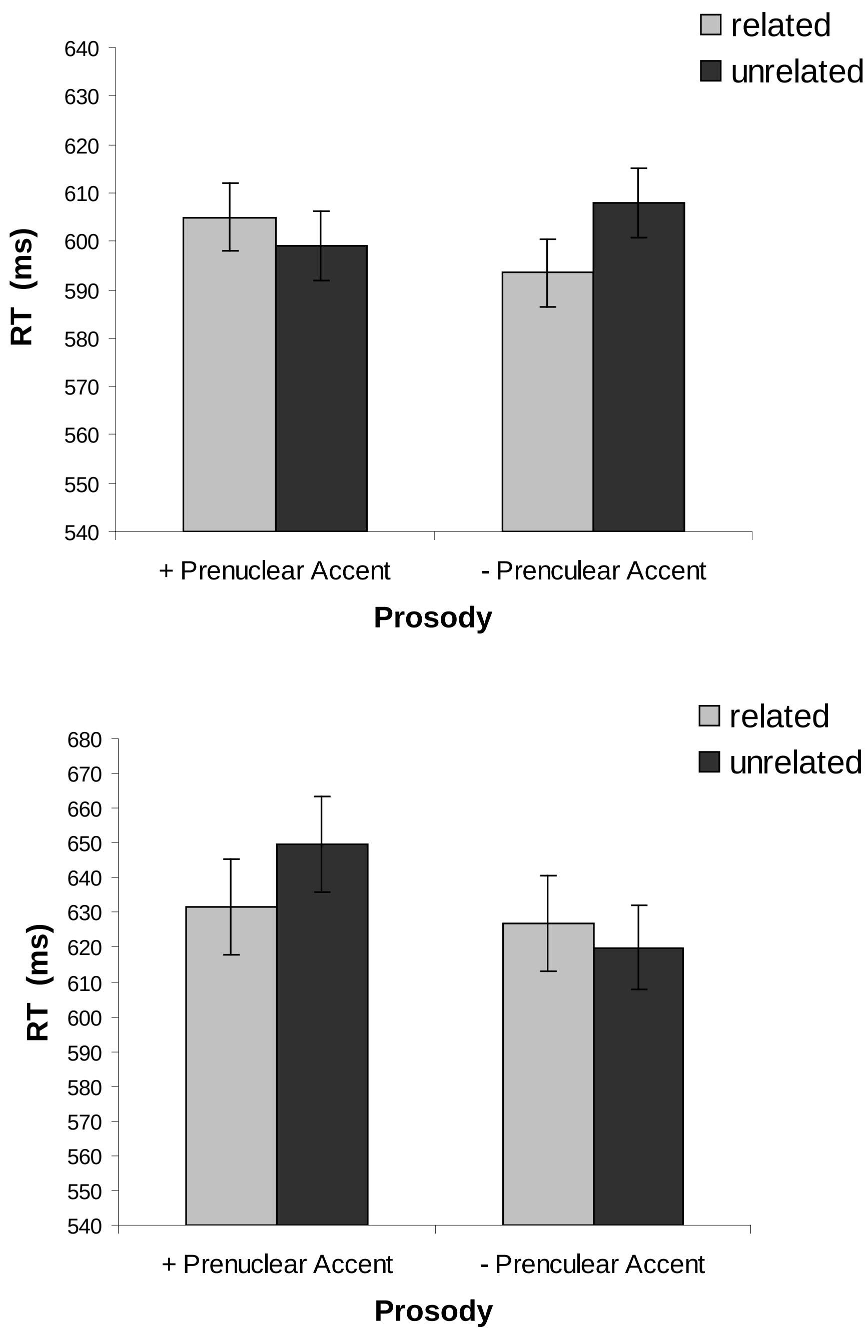

Figure 3. Mean reaction times for each of the experimental conditions in Experiment 2 for participants with low AQ-Comm scores (top) and those with high AQ-Comm scores (bottom). The “high AQ-Comm” group represents subjects with scores higher than one standard deviation above the group’s mean, the low group all others. Error bars show standard error.

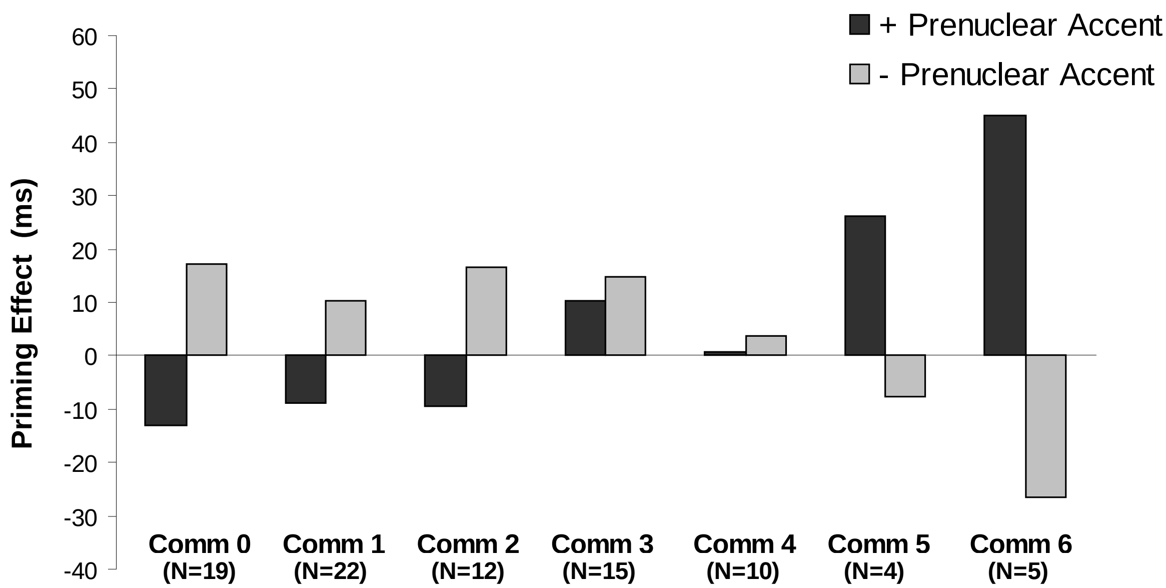

Figure. 4. Priming effects for the two prosodic conditions in Experiment 2, binned by participants' AQ-Comm score (ranging from 0 to 6). Priming effects are mean reaction time differences (in milliseconds) between responses to targets after unrelated control primes versus after related primes (positive numbers reflect facilitation relative to the control condition; negative numbers reflect inhibition). Higher AQ-Comm scores indicate more "autistic"-like communication skills. (Number of subjects in each bin is shown in parentheses).

can be seen, both facilitation of targets and inhibition of targets covaries with AQ-Comm; participants at the very low end of the spectrum (e.g., AQ-Comm=0-2) are very sensitive in the direction expected (based on production studies), showing robust facilitation of related targets in sentences with no nuclear accent, and inhibition when there is a prenuclear accent. In the mid-range of AQ-Comm scores, however, facilitation is less reliable, and there is no inhibition. Finally, at the very high end of the AQ-Comm spectrum (i.e., participants with very poor communication skills), we find a pattern that is the inverse relative to those on the low end of the spectrum. While it must be noted that the average differences are least reliable at this higher end of the range of scores (because the distribution is such that fewer participants scored in this region), the relationship is clear.

Finally, there was also a significant two-way interaction between prime type and AQ-Comm, indicating that the relatedness of the prime was a better predictor of priming overall for participants with higher AQ-Comm scores, regardless of prosody. There was also a marginally significant trend in the direction of high relative prominence on a prime being associated with faster reaction times overall, regardless of the relatedness between prime and target, or AQ-Comm score.

3. Discussion

3.1 The prosodic realization of focus structure

In the current study, cross modal priming was used in two experiments to explore a question that has been the topic of considerable interest in phonetic, phonological, and psycholinguistic literature. This question is to what extent the size of the focus constituent is related to prosodic structure in English SVO constructions. As described in the beginning of the paper, broad and narrow focus sentences are distinct, semantically in terms of the propositional content, and pragmatically in terms of what information is informative in the sentence. However, because the location of the nuclear accent does not vary as a function of this contrast in the constructions under consideration, there is a high degree of prosodic ambiguity. The question was to what extent listeners have expectations about how other aspects of sentence prosody, in particular, prenuclear accents, are relevant to the contrast. We can now provide some answers to this question, in the context of the literature that has been considered.

First, in Experiment 1, we tested whether a sentence with a prominent nuclear accent and no prenuclear accents was appropriate to express a sentence with focus on the object and with focus on the entire VP. If listeners expectations are for the patterns reported in production studies, they should have found this prosodic structure ideal for the expression of narrow object focus, but less so for broad VP focus. Although there was a trend in the direction of priming being numerically greater for narrow focus, this was not significant. Rather, the primary finding of Experiment 1 was that cross modal associative priming found an SVO sentence with a single pitch accent on the object to be equally appropriate in both VP and object focus contexts. This is in agreement with earlier off-line studies (Gussenhoven 1983, Welby 2003) and is a significant finding for the theoretical literature, as preeminent theories of prosody-information structure mapping predict this result. In particular, focus projection theories predict that a nuclear accent on the head of an internal argument can project focus to the larger phrase, and thus VP focus never requires the verb to bear a pitch accent (as long as there is an argument such as an object that is accented).

Independent of prosodic considerations, the results of Experiment 1 can also be viewed as providing evidence for the psychological reality of the discourse-dependent semantic difference between broad and narrow focus. The significant main effect for focus size indicates that listeners were able to make decisions faster to targets that followed narrowly focused primes, regardless of whether or not they were related to those primes. This is the pattern we might expect if broader focus constituents required more processing resources. It seems reasonable that, other things being equal, the set of possible alternatives to broader constituents such as a VP should be larger than the set of alternatives to a single object, simply by virtue of the combinatorial potential of verbs and arguments. Larger alternative sets should, in principle, impose a greater processing cost if they are actually activating lexical and conceptual representations. It may also be that processing larger chunks of new information simply requires more processing resources than do smaller chunks, regardless of any additional focus semantic value. Since priming was not disrupted in the VP focus condition, this difference is probably best regarded as one of these general mechanisms, rather than a violation of listeners' prosodic expectations.

In Experiment 2, however, the situation was quite different. In Experiment 2, listeners heard sentences with the same information structure, namely narrow object focus. This was a particularly interesting situation to test, in terms of both the phonetic and the theoretical literature on the prosody-information structure interface. As noted above, speakers are known to disambiguate both broad and narrow focus productions, but there is some evidence that they may do so more reliably for narrow focus. As mentioned, Breen et al. (2010) reported a tendency for speakers to produce narrow object focus sentences that were somewhat more accurately identified by statistical classification and human listeners compared with broad (sentence) focus sentences. Thus there is reason to believe that, in the minds of listeners, expectations about what constitutes a prototypical prosodic realization may be stronger for narrow object focus than broad VP focus. This is also a point at which the theoretical literature, at least with respect to focus projection, diverges. Gussenhoven's (1984/1999) model predicts a genuine ambiguity, since prenuclear accentuation is assigned to these structures by a late phonological process; Selkirk's (1995) model, however, predicts narrow object foci to be appropriate only without a prenuclear accent on the verb, since a prenuclear accent on the verb would incorrectly mark its discourse status as "new".

The results of Experiment 2 showed that priming was observed for narrow focus only if the sentence lacked a prenuclear accent. Overall, there was also a tendency towards inhibition of lexical decisions when narrow foci where presented with a prenuclear accent on the verb. These results are consistent with production patterns reported, and they are also consistent with the basic predictions of Selkirk's model. They are not consistent, however, with models that assume ambiguity (Gussenhoven 1999, Ladd 1996), including those that assign no information structural significance to accentuation in the prenuclear domain at all (Büring 2007). Overall, listeners were sensitive to prenuclear prominence in their interpretation of the sentence's information structure.

3.2 Individual Differences

An additional and important finding of the present study, however, is that it is not truly possible to speak about listeners "overall". Specifically, the results of Experiment 2 demonstrated very clearly that individual differences in listeners' working memory and autistic traits were important to predicting results. First, it was found that listeners with lower working memory were slower to respond. This may have been due to the fact that in Experiment 2, unlike in Experiment 1, the prosody of the test sentence varied from trial to trial. Having to integrate an unpredictable (and sometimes infelicitous) prosodic structure with the previous discourse context on each trial may have presented a greater burden on attentional resources, slowing down participants with lower working memory resources. This is consistent with recent findings from brain imaging; Kristensen, Wang, Petersson, and Hagoort (2012) report that, in addition to activating brain areas involved in semantic and pragmatic processing, intonational pitch accents are associated with heightened activation in the same region used by a spatial attention task.

Whereas working memory capacity exhibited rather general effects, listeners' autistic traits interacted crucially with the prosodic manipulation. However, the relation between AQ-Comm and priming patterns was more complex than simply distinguishing listeners who were sensitive to prosody and those who were not. As can be seen in priming results from Experiment 2 plotted in Figure 4, there are two separate processes at play, both of which are robustly related to AQ-Comm: facilitation and inhibition. Let us consider these two processes separately.

In the case of facilitation, there is in fact evidence that listeners with high AQ-Comm were less sensitive to prosody overall. The presence of a significant two-way interaction between prime type and AQ-Comm in the model from Experiment 2 indicates that poor communication skills were associated with more reliable priming of targets, to some extent regardless of other factors, including prosody. The consequence of this can be seen in Figure 4, where at the lowest end of the AQ-Comm spectrum (scores from 0-2), priming occurs only in the –prenuclear accent condition, but is less closely tied to prosody for those with higher scores. Additionally, at the very top of the AQ-comm continuum (scores 5-6), priming, when it occurs, is much stronger than for other listeners.

The pattern of results for listeners with poorer communication skills seems to resemble the more reliable and stronger priming effects found in studies where primes are presented in isolation rather than in sentences (Tabossi 1988, Norris et al. 2006). This likely reflects the "shallower" processing of simple lexical-semantic relationships, which is also thought to be typical of the processing of unfocused information (Sanford and Garrod 1998, Sanford and Sturt 2002). These results are therefore consistent with Nieuwland et al.'s (2010) findings that pragmatic "violations" only distinguished high and low AQ-Comm individuals if the violation required a focus-dependent interpretation. They are also consistent with similar findings reported by Xiang et al. (2012), who found AQ-Comm-dependent differences in the interpretation of sentences with "*only*", which is regarded as "focus sensitive", but not for "*every*", which may lack such focus-related meaning (Beaver and Clark 2003). Taken together, the robust facilitation found in the absence of what should have prosodically been the "effective context" (Foss and Ross 1983) suggests that listeners with poorer communication skills used neither prosody nor the discourse context to generate sentence-level representations in the task.

The relevance of sentence-level interpretation may also be important to understanding the patterns of inhibition observed. With respect to individual differences relating to AQ-comm, the most striking result in Experiment 2 was perhaps the inhibition of targets following primes that followed unaccented verbs. This indicates that high and low AQ-Comm individuals differed not just in their sensitivity to prosody, but in how they used it. While the details are difficult to deduce at this point, there may be at least two possible mechanisms for this inhibition. The first is that high AQ-Comm is associated with the activation of the "wrong" alternative set, and thus the suppression of the correct alternatives. This may be consistent with a recent proposal by Husband and Ferreira (2012), who claim that the generation of alternative sets is a two-step process that proceeds as follows. First, a contrastive interpretation for a word activates all semantically related words; this activation is then followed by a separate inhibition mechanism. According to such an account, individuals with high AQ-Comm differ from their low AQ-Comm counterparts in semantic and pragmatic interpretation – generating the "wrong" alternative set. They would not differ from low AQ-Comm listeners, however, in their basic use of prosody to generate that alternative set.

However, given the evidence that context-dependent semantic interpretations are exactly what individuals with autistic traits tend to underutilize, this scenario seems somewhat less likely. Another possibility is that inhibition is more related to limitations on processing resources. In fact, it has been argued that inhibition in associative priming can sometimes occur as the result of attentional suppression. For example, according to Marí-Beffa, Houghton, Estévez, and Fuentes (2000), inhibition tends to occur due to the need to suppress word-level representations in order to attend to the construction of a larger, sentence-level semantic interpretation. However, these authors, as well Tipper (1985), found that inhibition to targets also occurred when participants were made to actively ignore related primes. It may be that primes that were relatively highly prominent (i.e. those following unaccented verbs) placed more of a burden on attentional resources, and that high AQ-Comm individuals were most sensitive to this. As discussed above, there is evidence that the processing of prosodic prominence involves heavy use of such resources (Kristensen et al. 2012), and so it is possible that individuals with higher

AQ-Comm actively shift attention away from prosodic prominence. Clearly this matter requires further study.

3.3 Implications for the interpretation of prosodic prominence

Theories of the prosody-information structure interface regard prominence marking as largely paradigmatic. That is, prosodic prominence is marked by such categorical distinctions as the location of the nuclear pitch accent, or the presence versus absence of prenuclear accents, treating them as largely independent from one another. Additionally, they treat prenuclear accents as largely unrelated to information structure. Another possible conception of prominence marking, however, is one in which prenuclear and nuclear accents have a more syntagmatic relation to one another, and thus both are related to the expression of information structure in a fundamental way. While this has not been the standard theoretical approach, more syntagmatically structured prominence relations between intonational objects has, in different ways, been proposed previously (e.g., Ladd 1990, Dilley 2005, Calhoun 2010).

While the results of the experiments presented above are consistent with both paradigmatic and syntagmatic conceptions of prominence marking, I propose that the latter characterization is necessary to relate the perception and processing results to patterns observed in the production studies reviewed at the beginning of this paper. There it was found that, when gradient and holistic measures were considered, speakers were found to manipulate both prenuclear and nuclear material, often in tandem. If the prosodic encoding of information structure were primarily about the addition or subtraction of pitch accents, systematic behavior of this sort is highly unexpected. Since a syntagmatic model of prominence marking provides a better account for speakers' encoding of the focus size contrast, a fruitful task may therefore be to explore what other information structural contrasts are more insightfully analyzed this way. Indeed, recent research has pursued this for the theme/rheme (topic/focus) distinction (Calhoun 2010, Calhoun 2012). One challenge for such an approach, however, will be to develop an explicit account of how phonetic parameters are used by speakers to encode different kinds of contrasts. For example, if a relative prominence relation between the prenuclear and nuclear pitch accent is the phonological object, it remains underspecified to what extent listeners can substitute one phonetic cue to prominence (i.e., duration, intensity) for another (e.g., f0), and when. Such questions, however, rarely arise in models emphasizing accentuation as a more paradigmatic feature; exploring the implications of relative prominence marking of offers a promsing approach to the study of the prosodic encoding of information structure.

References

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): evidence from Asperger Syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders* 31, 5-17.

Bates, D. and Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. Available at: *http://CRAN.R-project.org/package=lme4*.

Baumann, S., Grice., Steindamm, S. (2006). Prosodic marking of focus domains: categorical or gradient? *Proceedings of Speech Prosody 2006*, 301-304.

Baumann, S., Becker, J., Grice, M., Mücke, D. (2008). Tonal and articulatory marking of focus in German. *Proc. of the XVIth International Congress of Phonetic Sciences*, 1029-1032.

Beaver, D. and Clark, B. (2003). Always and only: Why not all focus sensitive operators are alike. *Natural Language Semantics* 11, 323-362.

Birch, S. and Clifton, C. (1995). Focus, accent and argument structure: effects on language comprehension. *Language and Speech* 38, 365-391.

Bishop, J. (2012). Information structural expectations in the perception of prosodic prominence. In G. Elordieta and P. Prieto (Eds.): *Prosody and Meaning (Trends in Linguistics)*, 239-270. Berlin: Mouton de Gruyter.

Blutner, R. and Sommer, R. (1988). Sentence processing and lexical access: the influence of the focus identifying task. *Journal of Memory and Language* 27, 359-367.

Braun, B. and Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes* 25, 1024-1043.

Breen, M., Fedorenko E., Wagner., M., and Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes* 25, 1044-1098.

Büring, D. (2007). Intonation, semantics and information structure. In Gillian Ramchand and Charles Reiss (Eds.): *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press.

Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language* 86, 1-42.

Calhoun, S. (2012). The theme/rheme distinction: Accent type or relative prominence? *Journal of Phonetics* 40, 329-349.

Cole, J., Mo, Y., and Baek, S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes* 25, 1141-1177.

Daneman, M. and Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.

Dilley, L. (2005). *The phonetics and phonology of tonal systems*. Ph.D. Dissertation, MIT.

Dohn A., Garza-Villarreal E., Heaton P., and Vuust P. (2012). Do musicians with perfect pitch have more autism traits than musicians without perfect pitch? An empirical study. *PLoS ONE* 7(5): e37961. doi:10.1371/journal.pone.0037961.

Eady, S. and Cooper, W. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America* 80, 402-415.

Eady, S. and Cooper, W., Klouda, G., Mueller, P., and Lotts, D. (1986). Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments. *Language and Speech* 29, 233-251.

Fedorenko, E., Gibson, E., and Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. *Journal of Memory and Language* 54, 541-553.

Féry, C. and Krifka, M. (2008). Information structure Notional distinctions, ways of expression. In Piet van Sterkenburg (Eds.): *Unity and diversity of languages*, 123-136. Amsterdam: John Benjamins.

Foss, D. and Ross, J. (1983). Great expectations: context effects during sentence processing. In G. Flores D'Arcais and R. Jarvella (Eds.): *The process of language understanding*. Chichester: Wiley. 169-191.

Fowler, C. (2006). Compensation for coarticulation reflects gesture perception, not spectral contrast. *Perception and Psychophysics* 68, 161-177.

Ganong, W. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6, 110-125.

Grice, M., Baumann, S., and Benzmüller, R. (2005). German intonation in autosegmental-metrical phonology. In S.-A. Jun (Ed.): *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, 55-83.

Gussenhoven, C. (1983). Testing the reality of focus domains. *Language and Speech* 26, 61-80.

Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

Gussenhoven, C. (2005). Transcription of Dutch intonation. In S.-A. Jun (Ed.): *Prosodic typology: the phonology of intonation and phrasing*, 118-145. Oxford: Oxford University Press.

Hanssen, J., Peters, J., and Gussenhoven, C. (2008). Prosodic effects of focus in Dutch declaratives. *Proceedings of Speech Prosody 2008*, 609-612.

Hoekstra, R., Bartels, M., Cath, D., and Boomsma, D. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): A study in Dutch population and patient groups. Journal of Autism and Developmental Disorders 38, 1555-1566.

Husband, M. and Ferreira, F. (2012) Generating contrastive alternatives: Activation and suppression mechanisms. Paper presented at the CUNY Conference on Human Sentence Processing, The Graduate Center of the City University of New York.

Jun, S.-A. (2010). The Implicit Prosody Hypothesis and overt prosody in English. *Language and Cognitive Processes* 25, 1201-1233.

Kong, E.-J. and Edwards, J. (2011). Individual differences in speech perception: Evidence from visual analogue scaling and eye-tracking. *Proceedings of the XVIIth International Congress of Phonetic Sciences*, 1126-1129.

Kristensen, L., Wang, L., Petersson, K., and Hagoort, P. (2012). The interface between language and attention: Prosodic focus marking recruits a general attention network in spoken language comprehension. *Cerebral Cortex*. Advance online publication: doi:10.1093/cercor/bhs164.

Ladd, R. (1980). *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.

Ladd, R. (1990) Metrical representation of pitch register. In J. Kingston and M. Beckman (Eds.): *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge University Press.

Ladd, R. (1996). *Intonational Phonology*. [Cambridge Studies in Linguistics 79]. Cambridge, UK: Cambridge University Press.

Makashay, M. (2003). Individual differences in speech and non-speech perception of frequency and duration. Unpublished Ph.D. Dissertation, The Ohio State University.

Mann, V. (1980) Influence of preceding liquid on stop consonant perception. *Perception and Psychophysics* 28: 407-12.

Mann, V. and Repp, B. (1981). Influence of preceding fricative on stop consonant perception. *Journal of the Acoustical Society of America* 69, 548-558.

Marí-Beffa, P., Houghton, G., Estévez, A., and Fuentes, L. (2000). Word-based grouping affects the prime-task effect on semantic priming. *Journal of Experimental Psychology: Human Perception and Performance 26*, 469-479.

Mitterer, H. (2006). On the causes of compensation for coarticulation: Evidence for phonological mediation. *Perception and Psychophysics* 68, 1227–1240.

Nieuwland, M., Ditman, T., and Kuperberg, G. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language* 63, 324-346.

Norris, D., Cutler, A., McQueen, J., Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology* 53, 146-193.

R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics* 1, 75-116.

Ruta, L., Mazzone, D., Mazzone, L., Wheelwright, S., and Baron-Cohen, S. (2011). The Autism-Spectrum Quotient: Italian version: a cross-cultural confirmation of the broader autism phenotype. *Journal of Autism and Developmental Disorders* 42, 625-633.

Sanford, A. and Garrod, S. (1998). The role of scenario mapping in text comprehension. *Discourse Processes* 26, 159-190.

Sanford, A. and Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences* 6, 382-386.

Selkirk, E. (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass.: MIT Press.

Selkirk, E. (1995). Sentence prosody: intonation, stress, and phrasing. In J. Goldsmith (ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell, 550-569.

Sityaev, D. and House, J. (2003). Phonetic and phonological correlates of broad, narrow and contrastive focus in English. *Proceedings of the XVth International Congress of Phonetic Sciences*, 1819-1822.

Snedeker, J. and Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48: 103-130.

Sonié, S., Kassai, B., Pirat, E., Bain, P., Robinson, J., Gomot, M., Barthélémy, C., Charvet, D., Rochet, T., Tatou, M., Assouline, B., Cabrol, S., Chabane, N., Arnaud, V., Faure, P., and Manificat, S. (2012). The French version of the Autism-Spectrum Quotient in adolescents: A cross-cultural validation study. *Journal of Autism and Developmental Disorders*, 1-6 (online version accessed: Doi: 10.1007/s10803-012-1663-0).

Stewart, M. and Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition* 109, 157-162.

Surprenant, A., and Watson, C. (2001). Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *Journal of the Acoustical Society of America* 110, 2085-2095.

Swets B., Desmet T., Hambrick D., and Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: a psychometric approach. *Journal of Experimental Psychology: General* 136, 64-81.

Swinney, D. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 18, 645-659.

Tabossi, P. (1988). Effects of context on the immediate interpretation of unambiguous nouns. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14, 153-162.

Tabossi, P. (1996). Cross-modal semantic priming. *Language and Cognitive Processes* 11, 569-576.

Tipper, S. (1985). The negative priming effect: inhibitory effects of ignored primes. *Quarterly Journal of Experimental Psychology* 37A, 571-590.

Turk, A., Nakai, S., and Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In Stefan Sudhoff, Denisa Lenertova, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter, and Johannes Schließer (Eds.). *Methods in Empirical Prosody Research (Language, Context and Cognition 3)*. Berlin: Walter de Gruyter.

Unsworth, N., Heitz, R., Schrock, J., and Engle, R. (2005). An automated version of the operation span task. *Behavior Research Methods* 37, 498-505.

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., and Tojo, Y. (2006). The Autism-Spectrum Quotient (AQ) in Japan: A cross-cultural comparison. *Journal of Autism and Developmental Disorders* 36, 263-270.

Welby, P. (2003). Effects of pitch accent position, type, and status on focus projection. *Language and Speech* 46, 53-81.

Xiang., M., Giannakidou, A., and Grove, J. (2012). Two stages of NPI licensing: An ERP study. Paper presented at the CUNY Conference on Human Sentence Processing, The Graduate Center of the City University of New York.

Xu, Y. and Xu, C. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33, 159-197.

Yu, A. (2010). Perceptual compensation is correlated with individuals' "autistic" traits: implications for models of sound change. *PLoS ONE*, *5*, doi:10.1371/journal.pone.0011950.

Yu, A., Grove, J., Martinović, M., Sonderegger, M. (2011). Effects of working memory capacity and "autistic" traits on phonotactic effects in speech perception. *Proceedings of the XVIIth International Congress of Phonetic Sciences*, 2236-2239.

Appendix: Test Sentences, Primes and Targets

Item	Test Sentence with sentence-final prime	Control-prime	Visual Target
1.	He amused the butler.	bachelor	maid
2.	He borrowed a screwdriver.	newspaper	hammer
3.	She broke the table.	teapot	chair
4.	He called a doctor.	donor	nurse
5.	I dropped my wallet.	bible	keys
6.	He fixed the window.	cable	door
7.	I found a penny.	bottle	nickel
8.	He hurt his arm.	horse	leg
9.	She identified the artist.	burglar	painting
10.	She impressed the tutor.	locals	teacher
11.	He kissed a blonde.	guest	brunette
12.	He made a cake.	scarf	pie
13.	He observed a dolphin.	stars	whale
14.	He offended an employee.	attendant	boss
15.	She ordered a whiskey.	movie	beer
16.	He bought a motorcycle.	supermarket	car
17.	He requested a pillow.	toothbrush	blanket
18.	He smelt coffee.	smog	tea
19.	He stole a brownie.	dollar	cookie
20.	She took my fork.	hat	spoon
21.	He wanted a pancake	flashlight	waffle
22.	He washed the dog.	bike	cat
23.	He won an apple.	trophy	pear
24.	He loves spinach.	cities	broccoli
25.	He annoyed a girl.	bird	boy
26.	He borrowed the salt	tape	pepper
27.	I bought boots.	books	shoes
28.	He broke the toaster.	picture	oven
29.	He called his sister.	student	mother
30.	He finished his soup.	chores	salad
31.	I fixed the sink.	fence	toilet
32.	I found a pencil.	cell phone	pen

Appendix: Focus Contexts

Item	VP Focus Context	Object Focus Context
1.	What did Robert do at the party? Did he keep to himself?	Who did Robert amuse at the party? Did he amuse the children?
2.	What did John do when he was here? Did he chat?	What did John borrow when he was here? Did he borrow a book?
3.	What did Natasha do when she found out? Did she scream?	What did Natasha break when she found out? Did she break the dishes?
4.	What John you do when he found out? Did he cry?	Who did John call when he found out? Did he call his mother?
5.	What did you do? Did you trip and fall?	What did you drop? Did you drop your hat?
6.	What did John do when he came home? Did he sit around?	What did John fix when he came home? Did he fix the refrigerator?
7.	What did you do? Did you forget something?	What did you find? Did you find someone's keys?
8.	What did John do? Did he forget about today?	What did John hurt? Did he hurt a person?
9.	What did Natasha do when she arrived? Did she teach?	Who did Natasha identify when she arrived? Did she identify the owner?
10.	What did Natasha do? Did she give up?	Who did Natasha impress? Did she impress the visitor?
11.	What did Robert do after the party? Did he leave?	Who did Robert kiss after the party? Did he kiss Mary?
12.	What did Robert do? Did he clean?	What did Robert make? Did he make something healthy?
13.	What did Robert do while he was there? Did he walk around?	What did Robert observe while he was there? Did he observe people?
14.	What did John do at the restaurant? Did he behave?	Who did John offend at the restaurant? Did he offend your friend?
15.	What did Natasha do? Did she get up and leave?	What did Natasha order? Did she order a pizza?
16.	What did Robert do with his money? Did he invest?	What did Robert buy with his money? Did he buy a Rolex?
17.	What did the executive do when he called up? Did he complain?	What did the executive request when he called up? Did he request a refund?
18.	When the executive came downstairs, what did he do? Did he read?	When the executive came downstairs, what did he smell? Did he smell smoke?
19.	What did John do when he came in? Did he clean up?	What did John steal when he came in? Did he steal the remote?
20.	What did Natasha just do? Did she ask for something?	What did Natasha just take? Did she take your seat?
21.	What did John do when you woke him up? Did he go back to sleep?	What did John want when you woke him up? Did he want a glass of water?
22.	What did John do this afternoon? Did he paint?	What did John wash this afternoon? Did he wash the car?
23.	What did the student do? Did he give up?	What did the student win? Did he win a ribbon?
24.	What did you hear about the actor? Does he work out?	What does the actor love? Does he love exercise?
25.	What did John do at the Anderson's party? Did he dance?	Who did John annoy at the Anderson's party? Did he annoy Mr. Anderson?
26.	What did Robert do when he came by? Did he snoop around?	What did Robert borrow when he came by? Did he borrow the broom?
27.	What did you do this afternoon? Did you work?	What did you buy this afternoon? Did you buy groceries?
28.	What did John do? Did he help out?	What did John break? Did he break a mirror?
29.	What did Robert do? Did he go to bed?	Who did Robert call? Did he call his friend?
30.	What did Robert do before he left the house? Did he clean up?	What did Robert finish before he left the house? Did he finish his homework?
31.	What did you do all day? Did you sleep in?	What were you fixing all day? Did you fix the roof?
32.	What did just do? Did you trip on something?	What did you find? Did you find a dollar