Prosody and meaning: On the production, perception and interpretation of prosodically realized focus

Judith Tonhauser
The Ohio State University

December 24, 2017

Commissioned article for Cummins, C. and N. Katsos (eds.) *Handbook of Experimental Semantics and Pragmatics*, Oxford: Oxford University Press

Abstract

The prosody of an utterance plays a significant role in determining the meaning of the utterance. Studying the contributions of prosody to meaning is complicated by several factors: i) prosody has multiple components in the speech signal, some with continuous expression, ii) utterances with the same meaning can differ in their prosodic realizations, and iii) there is cross-linguistic prosodic variation. Concentrating on information-structural focus, this article illustrates how experimental investigations advance our understanding of the intricate relationship between prosody and meaning. The article discusses how focus is prosodically realized in different languages, how listeners perceive and interpret prosodically realized focus and how prosodically realized focus interacts with contextual information about focus. Different methods used to explore prosodically realized focus and its perception and interpretation are covered. The article concludes by considering research on the prosody of semantic/pragmatic phenomena related to focus, such as contrastive topic and presupposition.

1 Introduction

The prosody of an utterance, which includes the grouping of the words of the utterance, the relative prominences of the words and the tonal movements, makes tremendous and diverse contributions to the meaning of the utterance. Prosody can convey the emotional status of the speaker, such as anger, disgust, fear or sadness (e.g., Brosch et al. 2010). Prosody distinguishes between speech acts: for instance, the American English declarative sentence *It's raining* is a request for confirmation when uttered with a rise at the end, but an assertion when uttered with a fall at the end (e.g., Gunlogson 2002). Prosody can also influence whether an interrogative sentence like *Do you want to do the laundry?* is understood as seeking information, as an invitation or as a request, i.e., the illocutionary act, and how annoyed, authoritative and polite the speaker is taken to be, i.e., the perlocutionary effect (Jeong and Potts 2016). Prosody can even influence the truth conditions of utterances. Whether, for example, the relative clause *who was on the balcony* in utterances of *Someone shot the servant of the actress who was on the balcony* is taken to modify *the actress* or *the servant* depends on the prosody of the utterance (e.g., Speer et al. 2011). Prosody also influences the truth conditions of utterances with focus sensitive operators like *only* (e.g., Jacobs 1983, von Stechow 1990, Rooth 1992, Beaver and Clark 2008). Thus, an utterance of *Kim only serves Sandy Johnnie Walker* in which *Johnnie Walker* is the prosodically most prominent expression is taken to mean that Kim serves Sandy nothing but Johnnie Walker, whereas an utterance in which *Sandy* is the prosodically most prominent expression is taken to mean that Kim serves nobody but Sandy Johnnie Walker.

This article examines the contributions of prosody to meaning by considering experimental research on a category of meaning that has been central to research on prosody and meaning, namely information-

structural focus. The study of information structure is concerned with the question of how the (morphological, syntactic, semantic, prosodic, etc.) properties of an utterance reflect the organization of the discourse in which the utterance occurs (see, e.g., Lambrecht 1994, Krifka 2006, Féry and Krifka 2008). In this article, 'focus' refers to an information-structural property of parts of utterances and is defined on the basis of questions: an expression is a focus of an utterance if the content of that expression answers the (possibly implicit) question that the utterance addresses (see, e.g., Roberts 2012/1998, Beaver and Clark 2008). In (1), for instance, where B's utterance addresses A's interrogative utterance, the proper name *Mary* answers the question denoted by the interrogative utterance and hence is the focus of B's utterance. In (2), the focus of B's utterance is the indefinite noun phrase *a tornado*.[1] In this article, brackets subscripted with 'F' ($[\ ]_F$) identify focused expressions. This notation should not be taken to convey any commitments about the syntactic representation of focus (but see, e.g., Rooth 1992). To avoid confusion, the term 'focus' is not used in this article to refer to a phonetic or phonological property of focused expressions (e.g., 'prosodic focus').

(1) A: Who saw the tornado?
B: $[\text{Mary}]_F$ saw the tornado. (adapted from Roberts 2012:1910)

(2) A: What did Mary see?
B: Mary saw $[\text{a tornado}]_F$.

Question-answer pairs like (1) and (2) are the main diagnostic for information-structural focus, regardless of whether focus is marked prosodically, morphologically or syntactically in a given language.[2] As these examples illustrate, foci can be expressions whose denotations are given, like the proper name *Mary* in (1), or new, like the indefinite noun phrase *a tornado* in (2). Some works, like e.g., Katz and Selkirk 2011, reserve the term 'focus' for contrastively focused expressions, to the exclusion of expressions providing new information, as discussed below.

This article proceeds as follows. Section 2 provides a brief introduction to prosody and how experimental investigations can meet challenges to exploring the contributions of prosody to meaning. Section 3 examines how expressions that are information-structural foci are prosodically realized, and section 4 how prosodically realized foci are perceived and interpreted. Both sections 3 and 4 discuss data from a variety of languages, the influence of context, how focus size is distinguished and the role of experimental methods in exploring the prosody of focus. The article concludes in section 5 with a brief discussion of the contributions of prosody to semantic/pragmatic phenomena related to focus, such as contrastive topic and presupposition. Additional linguistic phenomena influenced by focus are discussed in Kim's chapter on focus in this volume.

2 Prosody: A brief primer

This section is a brief introduction to prosody and to challenges to exploring prosody and meaning. For more detailed introductions to prosody from different perspectives see, e.g., Hirst and Di Cristo 1998:ch1, Pierrehumbert and Hirschberg 1990, Gussenhoven 2002, Grice and Baumann 2008, Ladd 2008 and Beckman and Venditti 2011.

[1] What is referred to as focus here is referred to as 'rheme' in other parts of the literature (e.g., Halliday 1985, Roberts 2012). Other authors define 'focus' on the basis of sets of alternatives (e.g., Rooth 1992), givenness (e.g., Schwarzschild 1999) or a combination of givenness and providing an answer to the question under discussion (e.g., Büring 2007). Different definitions of 'focus' are discussed in Kim's chapter in this volume

[2] Some languages have been reported to not use prosody to mark information-structural focus, including Wolof (Rialland and Robert 2001), Yukatek (Kügler and Skopeteas 2006, 2007), Chichewa, Chitumbuka and Durban (Downing 2008), Nɬeʔkepmxcin (Koch 2008), Northern Sotho (Zerbian 2006) and K'iche' (Yasavul 2013, Burdin et al. 2015). For discussion of prosodic, morphological and syntactic marking of focus see, e.g., Féry and Krifka 2008 and Zimmermann and Onea 2011. For the hypothesis that information-structural focus is marked prosodically in all languages see Roberts 1998.

Any utterance has phonetic properties that can vary independently of the expressions uttered: for instance, in English, the sentence *Mary saw a tornado* can be uttered with prosodic prominence on *tornado* without changing the contribution the expression makes to the truth conditions of the sentence. Prosodic prominence on *tornado* can therefore be used to convey (what is referred to as) a post-lexical meaning. Cross-linguistically, the phonetic properties of utterances that can vary independently of the expressions uttered in order to convey post-lexical meanings include fundamental frequency, duration (of syllables or pauses) and intensity. Languages differ in whether a particular phonetic property is used to convey post-lexical meanings: for instance, a rise in fundamental frequency that is associated with the stressed syllable of an expression is a correlate of a post-lexical (high tone) pitch accent in English (e.g., *tornado* means 'tornado' regardless of whether it is realized with a high or a low tone pitch accent), but such a rise may be a correlate of a lexical pitch accent in Swedish (e.g., *anden* means 'the duck' or 'the spirit', depending on the pitch accent with which it is realized) or of tone in Mandarin Chinese (where, famously, *ma* means 'mother' or 'hemp' depending on whether it is realized with a high tone or a rising tone, respectively). Languages also differ in the phonetic properties that have become conventionalized bearers of post-lexical meaning and are thus part of the phonology of the language. The phonetic and phonological properties of a language that are used to convey post-lexical meaning constitute the prosody of the language.

The prosody of a language can provide cues about the organization of the utterance and its relation to the context in which it occurs: these cues include phrasing (how the expressions of the utterance are grouped), prosodic prominence (the relative prominence of the expressions in the utterance) and tonal movements (the tune of the utterance). Examples of such information cued by prosody include (see section 1) the attitude of the speaker, the speech act of the utterance, the syntactic and semantic relationships among the expressions of the utterance and the information structure of the utterance. In English, for instance, whether the relative clause *who was on the balcony* modifies *the servant* or *the actress* can be indicated by marking either one of the two noun phrases as the right edge of a phrase, by a combination of lengthening and hyperarticulation of segments preceding the edge, pitch movements associated with the edge and a pause (for detailed discussion see, e.g., Speer et al. 2011 and Wagner and Watson 2010). Thus, the study of prosody (or 'intonation', as it is also sometimes called) is not concerned with phonetic or phonological properties of an utterance that determine the lexical meanings of the expressions uttered, but with those properties that convey other types of meaning, under a broad use of the term 'meaning'. For an introduction to the autosegmental-metrical framework of intonational phonology assumed by many of the works reviewed in this article see, e.g., Beckman 1996, Ladd 2008 and Shattuck-Hufnagel and Turk 1996.

Given the contributions that prosody makes to natural language meaning, theories of meaning must capture how speakers use prosody to convey meaning and how listeners perceive and interpret the prosody of the utterances they hear. Studying the contributions that prosody makes to meaning is challenging for a number of reasons. First, since not all phonetic and phonological properties of an utterance contribute to post-lexical meaning, research on prosody must tease apart those that do from those that don't. This task is complicated by the fact that the prosody of any language is complex, with multiple components, some of which have continuous expression in the speech signal (see, e.g., Wagner and Watson 2010, Watson 2010 on prominence in production and comprehension). Second, since the vocal folds and speech patterns of the speakers of a language vary along a number of dimensions, the phonetic properties that underlie the prosody of the language do not have any absolute value for fundamental frequency, duration or intensity in the language, but vary by age, sex, etc. (e.g., men's speech is typically characterized by lower fundamental frequency values than women's). There are thus no invariant phonological correlates of information-structural notions such as focus (see, e.g., Féry 2007) and even the prosody of utterances that convey the same meaning is highly variable, both within and across speakers of the same language (e.g., Baumann et al. 2006, Ito and Speer 2006, Speer et al. 2011). In short, the mapping between the prosody of an utterance in any given language and the meaning of the utterance is not straightforward. Furthermore, there is to date no agreement about the nature of the interface between prosody and meaning (for discussion see, e.g., Prieto 2015). Open

questions include, for example, whether the interface is direct, or mediated by syntax (for a probabilistic perspective see, e.g., Calhoun 2010), and whether individual tonal movements or combinations thereof are the basic meaning bearing units of prosody (see, e.g., Gussenhoven 2002 for discussion and relevant references). And, finally, as noted above, there is cross-linguistic variation in the prosodic systems of languages, even closely related ones (see, e.g., Hirst and Di Cristo 1998, Jun 2005, 2014).

Experimental investigations can help meet these challenges to studying the contributions of prosody to meaning. Perception and comprehension experiments can, for instance, help tease apart those phonetic and phonological properties of utterances that influence post-lexical meaning from those that don't. Production experiments designed to explore the prosodic realization of utterances with particular meanings can serve to identify speaker-invariant prosodic cues to those meanings. Finally, cross-linguistic comparison is facilitated by running the same experiments with speakers of different languages. This article illustrates advances made by experimental investigations into the contributions of prosody to meaning, in particular the information structural notion of focus.

3 The production of prosodically realized focus

This section is concerned with the production of prosodically realized focus. Section 3.1 introduces prosodic cues to focus across languages, and argues that the prosody of focus in a given language may be independent of the typological characteristics of the prosodic system of the language. Section 3.2 shows that contextual information about focus influences the prosodic realization of focus, which has consequences for the question of whether the grammars of languages distinguish different types of focus, as discussed in section 3.3. The idea that differences in focus size are prosodically distinguished is motivated in section 3.4. Section 3.5 concludes with a discussion of methods used in research that investigates how focus is prosodically realized.

3.1 Prosodic realization of information-structural focus across languages

In languages that mark information-structural focus prosodically, focused expressions are generally more prosodically prominent than expressions that are not focused. Languages differ in the phonetic and phonological means to implement prosodic prominence. These means include the presence and type of pitch accents and boundary tones, duration of segments and pauses, fundamental frequency range, and the alignment of pitch targets to the segmental string (e.g., Jun 2005, Ladd 2008, Watson 2010, Féry 2013). In English, for instance, focused expressions are made relatively more prosodically prominent than non-focused expressions in the same phrase by realizing the focused expression with a pitch accent, a longer duration, an expanded pitch range and greater intensity (e.g., Cooper et al. 1985, Eady and Cooper 1986, Eady et al. 1986, Terken and Hirschberg 1994, Xu and Xu 2005, Breen et al. 2010). In Jun's (2005) typology of prosody, English is a head-prominence language since prosodic phrase-level prominence is marked on the phrase head, which is identified as the head of the phrase through pitch accenting. In Paraguayan Guaraní (Tupí-Guaraní), also a head-prominence language, focused expressions are distinguished from non-focused ones by the type of pitch accent realized, by the steepness of the slope of rising pitch accents and by duration (Clopper and Tonhauser 2011, 2013, Burdin et al. 2015, Turnbull et al. 2015).

Korean, unlike English and Paraguayan Guaraní, is an edge-prominence language, which means that phrase-level prosodic prominence is marked on the left or right edge of the prosodic phrase through boundary tones (the left edge, in the case of Korean; see Jun 2000, Kim 2004). A Korean utterance is typically phrased into accentual phrases and (larger) intonation phrases. An accentual phrase, which is slightly larger than a word, is marked by a particular tonal pattern. A focused expression is made relatively more prosodically prominent than non-focused expressions by expanding its pitch range, lengthening its initial segment and by reducing the pitch range on the following expressions and the duration of the expression before and after

the focused expression. Optionally, a boundary tone may be inserted before the focused expression and the following words may be dephrased, i.e., realized in the same phrase as the focused word (e.g., Jun 1993, Jun and Lee 1998, Jun and Kim 2008).

Jun (2005:441) proposed that the way in which information structure, including focus, is prosodically realized in a given language may depend to some extent on the prosodic structure of the language. This proposal was investigated on the basis of a study of the prosody of noun, adjective and noun phrase focus in American English, Paraguayan Guaraní, K'iche' (Mayan) and Moroccan Arabic in Burdin et al. 2015. In contrast to American English and Paraguayan Guaraní, which are head-prominence languages, Moroccan Arabic and K'iche' were described in Burdin et al. 2015 as head/edge-prominence languages, which means that they exhibit both head-prominence and edge-prominence marking (see Jun 2014). Burdin and her colleagues found that, contrary to Jun's proposal, "the prosodic realization of focus is orthogonal to the head-, head/edge-, and edge-prominence distinction" (p.262): for instance, focused expressions were lengthened in American English, Paraguayan Guaraní and Moroccan Arabic, but not in K'iche'. (See Burdin et al. 2015:§2.5 for discussion of other literature that also suggests this conclusion.) In sum, languages that mark information-structural focus prosodically rely on a range of phonetic and phonological cues, and which cues a language relies on may be independent of the typological characteristica of the language.

3.2 The prosodic realization of focus is context-dependent

Question-answer pairs are fundamental in research on information-structural focus: since interrogative utterances determine the foci of their answers, question-answer pairs allow researchers to explore how foci are (prosodically, morphologically or syntactically) realized. It has long been observed that answers must be congruent with the interrogative utterances they address (e.g., Paul 1880, 1919, von Stechow 1990, Rooth 1992), as shown on the basis of the question-answer pairs in (3): J's utterance in (3a), with *Turkish* prosodically prominent (as indicated by the capital letters), is congruent with M's question in (3a), i.e., judged to be acceptable in response to that question, but not congruent with M's question in (3b), i.e., judged to be unacceptable in response to that question. The reverse pattern holds for J's utterance in (3b), in which *David* is prosodically prominent.

(3) Mandy, Craige and Judith are having lunch at a place that serves Turkish, Lebanese and Irish coffee.

a. M: What kind of coffee does David like?
   J: David likes [TURkish]$_F$ coffee.

b. M: Who likes Turkish coffee?
   J: [DAvid]$_F$ likes Turkish coffee. (adapted from Tonhauser 2016:951)

It follows from question-answer congruence that the focus marking of an utterance constrains the question addressed by the utterance, regardless of whether the question is realized by an interrogative utterance, as in the examples in (3), or whether the question is implicit, as in many naturally occurring examples (see, e.g., Roberts 2012/1998, Beaver et al. 2017). The fact that what is focused in an utterance can be conveyed both by context (e.g., the interrogative utterances in (3)) and by prosody suggests the information-theoretic hypothesis that the prosodic realization of focus is enhanced when the context does not identify which expression is focused. This hypothesis was investigated by Turnbull et al. (2015): in their production experiments, speakers of American English and Paraguayan Guaraní played an interactive game in which they instructed a confederate (of the same language) to place a series of tiles depicting different objects into numbered boxes on a game board. The speakers' utterances were either made in a context in which it was predictable from the other game pieces (visible to the speaker and the confederate) whether the adjective, the noun or the entire noun phrase was focused, or in a context in which the focus was unpredictable from the other game pieces. For instance, when the English utterance *Put the orange flower in box two* was made in

a context in which the other game pieces were flowers of a variety of colors, adjective focus was predictable from the visual context; when the same utterance was made in a context in which the other game pieces consisted of a variety of objects, including flowers, in a variety of colors, adjective focus was not predictable from the visual context.

Turnbull and his colleagues found that the prosodic realization of focus was enhanced in both languages when the visual context did not predict which expression was focused. For instance, when adjective focus was not predictable from context, focused English adjectives were longer, realized with higher f0 peaks and more likely to be realized with a rising pitch accent than when adjective focus was predictable from context. In Paraguayan Guaraní, whether focus was predictable from context did not affect pitch accent type or duration, but the f0 slope of accented adjectives was steeper (regardless of whether the adjective was focused) when focus was not predictable from context than when it was. Thus, the prosodic realization of focus is affected in both languages by the contextual predictability of focus, but the influence of contextual predictability on prosodic realization of focus appears to be language-specific.

Context-dependence in the prosodic realization of focus was also illustrated in Beaver et al.'s (2007) study of second occurrence focus, a repeated focus: In B's utterance in (4), *vegetables* associates with the focus sensitive operator *only*, but *vegetables* is given. Second occurrence focus is marked with brackets subscripted with 'SOF'.

(4) A: Everyone already knew that Mary only eats [vegetables]$_{\text{F}}$.
B: If even [Paul]$_{\text{F}}$ knew that Mary only eats [vegetables]$_{\text{SOF}}$, then he should have suggested a different restaurant. (Beaver et al. 2007:247)

Beaver et al. (2007) found that second occurrence foci are not realized with a pitch accent, but distinguished from unfocused expressions by a longer duration and a higher energy. Thus, whether a focus is realized in American English with a pitch accent depends on the context in which the focus occurs.

In sum, the prosodic realization of focus in a given language depends on the contextual cues to focus and the context in which the focused expression is realized.

3.3 Prosodic realization of other information-structural properties of foci

Focused expressions denote answers to the (explicit or implicit) questions addressed by the utterances in which they occur, but they may differ along other information-structural dimensions. In the examples in (5), for instance, the noun phrase *a coffee* is the focus of A's utterances, but the denotation of the noun phrase differs along other information-structural dimensions: the denotation is new information in (5a), it is previously mentioned and contrasted with the other member of a contextually-given set, namely tea, in (5b), and it is used to correct an interlocutor's assumption in (5c).

(5) a. Q: What would you like to drink?
A: I'll have [a coffee]$_{\text{F}}$, please.
b. Q: Would you like to have a coffee or a tea?
A: I'll have [a coffee]$_{\text{F}}$, please.
c. At the breakfast buffet, A may have coffee or tea, and has just been given a cup of tea.
A: I don't want tea, I want [a coffee]$_{\text{F}}$.

Given that focus is not the only meaning conveyed prosodically (see section 1), it is unsurprising that foci that differ in their information-structural properties may differ in their prosodic realization. The three studies discussed in the following illustrate this for foci that differ along the lines of the examples in (5).

In Ito et al. 2004, American English-speaking participants gave instructions to a confederate about how to decorate a tree, thereby producing sequences of utterances consisting of a color adjective and an object-denoting noun, such as ... *blue ball* ... *blue house*. The noun of an adjective-noun expression was defined

as contrastive if it was immediately preceded by an adjective-noun expression with the same adjective but a different noun, like *house* in . . . *blue ball* . . . *blue house* (likewise for adjectives). They found that adjectives and nouns whose denotations were contrastive (whether new or given) were more likely to be realized with a L+H* pitch accent[3] than adjectives and nouns whose denotations were (non-contrastive and) new.

In Breen et al.'s (2010) experiments, speakers uttered answers and listeners had to identify the question they were answering: e.g., an utterance of *DAmon fried an omelet this morning* answers *Who fried an omelet this morning?* but not *What did Damon fry this morning?* (recall that capital letters indicate prosodic prominence). An expression was defined as contrastive in the presence of an explicit contrast set: while *Damon* in the just-mentioned congruent question-answer pair is not contrastive, it is taken to be contrastive when *DAmon fried an omelet this morning* is uttered in response to *Did Harry fry an omelet this morning?*. In an experiment in which speakers were given feedback about whether listeners were successful in identifying the question they were answering, speakers' utterances distinguished between contrastive and non-contrastive subjects, verbs and objects, with contrastive expressions realized with longer durations and higher intensity than (given or new) non-contrastive ones. Surprisingly, non-contrastively focused words were produced with higher f0 than contrastively focused ones, a finding which may be in contrast to that of Ito et al. 2004, given that L+H* pitch accents are typically realized with a higher f0 than H* pitch accents (for discussion and references see, e.g., Ito and Speer 2008).

Katz and Selkirk (2011) also found differences in the prosodic realization of contrastive and (non-contrastive) discourse new expressions. Their production experiment compared the realization of expressions like *that Modigliani* and *MoMA* in utterances like *So he would (only) offer that Modigliani to MoMA* in which both expressions were discourse new, or only one was discourse new and the other one was contrastive. For Katz and Selkirk (2011), an expression was contrastive if it associated with a focus sensitive expression like *only*. They found that contrastively focused expressions did not differ from discourse new ones in terms of pitch accenting or phonological phrasing, but that contrastively focused constituents had a longer duration, a greater relative intensity and a greater f0 movement than discourse new ones in the same position. (For relevant studies in other languages see, e.g., Krahmer and Swerts 2001 on Dutch, D'Imperio 2001 on Neapolitan Italian, and Baumann et al. 2006 and Braun 2006 on German.)

These three studies suggest that prosody does not merely distinguish focused and non-focused expressions (section 3.1), but also distinguishes whether a focused expression denotes new or contrastive information. It is important to note, however, that the term 'contrastive' was defined differently in the three studies reviewed here: its definition involved an expression in the immediately preceding utterance (Ito et al. 2004), a question that the utterance was correcting (Breen et al. 2010), and association with a focus sensitive expression (Katz and Selkirk 2011). The fact that the studies explored three different (though possibly related) notions of 'contrastive' focus may very well account for the differences observed in the three studies in how prosody distinguishes contrastively and non-contrastively focused expressions. These studies thus highlight the importance of unambiguously defining the meaning categories whose prosodic realizations are explored. (For discussion see also Kim's chapter on focus in this volume.)

The findings of the three studies also bear on the question of whether foci that differ in other information-structural properties are grammatically distinguished. Katz and Selkirk (2011), for instance, argued that contrastive foci and discourse new foci are represented differently in the syntax and, thereby, receive distinct interpretations. (For related proposals see, e.g., Kiss 1998, Vallduvi and Vilkuna 1998, Selkirk 2007.) An alternative position (which is compatible with foci being syntactically represented or not) is one on which differences between foci are attributable to context and, hence, need not be grammatically represented (see, e.g., Rooth 1992, Roberts 1998, Büring 2007).

[3]The L+H* pitch accent is a high tone pitch accent that is preceded by a low tone. For an introduction to the Tones and Break Indices (ToBI) annotation system see Beckman and Ayers 1997.

3.4 Prosodic marking of focus size

This section shows that prosody provides cues to the size of the focus of a given utterance. B's utterances of *Kim baked a cake* in (6a-c), for instance, differ in whether the focus is the direct object, the verb phrase or the entire sentence, respectively.

(6) a. A: What did Kim bake today?
B: Kim baked [a cake]$_F$.
b. A: What did Kim do today?
B: Kim [baked a cake]$_F$.
c. A: What happened?
B: [Kim baked a cake]$_F$.

Since B's utterances in (6a-c) can all be realized with pitch accents on *cake*, and no other pitch accents, some authors assume that prosodic marking of focus is systematically ambiguous with respect to whether the direct object, the verb phrase or the entire clause is focused (e.g., Gussenhoven 1983, Selkirk 1996). Accordingly, Selkirk's (1996) widely adopted Focus Projection Hypothesis predicts that a pitch accented direct object can 'project' focus from the direct object to the verb phrase and to the entire utterance. (See Ladd 2008:ch.6 for a discussion of focus projection.)

However, experimental research on the prosodic realization of focus has provided empirical evidence that the size of focus can be prosodically distinguished. For American English, Breen et al. (2010) found that speakers produce subject-verb-object sentences differently depending on whether the entire sentence or only the object is focused. Specifically, they found that when only the object is focused, the object is produced with the highest maximum f0, the longest duration and the maximum intensity compared to other words in the sentence, whereas the words of the sentence were more prosodically similar to one another when the entire sentence was focused. Baumann et al. (2006) investigated the prosodic realization of German sentences with sentence focus, verb phrase focus and object focus, and found that the narrower the focus was, the more likely the direct object was to be produced with a longer duration and with a higher f0 peak relative to previous f0 peaks, and the less likely the verb was to be realized with an accent. For European Portuguese, Frota (2002) found that the object is realized with a longer duration and a different pitch accent when only the object is focused than when the entire sentence is focused. Similarly, Hanssen et al. (2008) showed that, in Dutch, sentence and object focus are distinguished by the duration of the object and by the timing, scaling and slope of f0 targets and movements.

Focus size is also distinguished in languages genetically unrelated to European ones. For Korean, for instance, a verb-final language, Kim et al. (2006) found that verb phrase focus was distinguished from sentence focus by inserting an intonational phrase boundary, by raising the pitch peak of the first word of the verb phrase and by lengthening the first and second words in the verb phrase. Jun and Kim (2008) observed that the object was relatively more prominent in object focus than in verb phrase focus: in object focus, the f0 peak of the object was higher and the object was longer than in verb phrase focus, and the verb phrase was more likely to be dephrased or produced with a reduced pitch range in object than in verb phrase focus.

Finally, focus size is also marked within the noun phrase. As mentioned above, Burdin et al. (2015) compared the prosodic realization of focus in utterances in which the adjective, the noun or the entire noun phrase was focused across four genetically unrelated languages. For American English, they found that noun focus is distinguished from noun phrase focus in that the adjective is less likely to be unaccented when the noun phrase is focused than when the noun is focused. (For a similar finding in Dutch see Krahmer and Swerts 2001.) For Paraguayan Guaraní, in which the noun precedes the adjective, they found that adjectives were longer when the noun phrase was focused than when the adjective was focused. In Moroccan Arabic, the noun again precedes the adjective. Here, Burdin and her colleagues found that both the adjective and the

noun were longer when the noun phrase was focused than when the adjective was focused. (K'iche' did not show any prosodic marking of focus; see footnote 2.)

In sum, genetically and typologically unrelated languages use a variety of prosodic means to lend prominence to particular parts of an utterance and to thereby identify the size of the focus of the utterance. These findings suggest that the assumption that underlies Selkirk's Focus Projection Hypothesis, that prosodic focus marking does not distinguish focus size, may not be empirically adequate, at least from the production perspective. Section 4.4 addresses whether listeners attend to prosodic cues to focus size.

3.5 Experimental tasks to explore the prosodic realization of focus

Experimental research on the prosodic realization of focus relies on a variety of tasks: recordings are made of individual participants reading scripts (e.g., Beaver et al. 2007, Katz and Selkirk 2011), of pairs of participants reading scripts to one another (e.g., Clopper and Tonhauser 2013), of participants completing a task with a confederate (e.g., Ito and Speer 2008, Burdin et al. 2015) and of pairs of participants engaging in an interactive task (e.g., Watson et al. 2008, Breen et al. 2010). Compared to empirical claims about the prosodic realization of focus that are based on individual researchers' intuitions, experimental investigations have the advantage of allowing researchers to quantify phonetic and phonological cues to focus, and to distinguish between speaker-invariant and speaker-specific prosodic cues to focus.

The aforementioned tasks differ significantly in the extent to which the elicited utterances resemble conversational speech, with participants reading scripts being the least and participants engaging in interactive tasks being the most like conversational speech. An advantage of scripted language is that researchers have full control over the content, length and structure of the produced utterances, thus facilitating the comparison of productions of pairs of utterances that differ only in the variables of interest. However, as discussed in Ito and Speer 2006, one of the worries of investigating prosody on the basis of productions of read, scripted language is that since such productions differ from conversational speech, the results of the investigation have limited generalizability. Ito and Speer therefore urge researchers to rely on tasks that examine "the relationship between produced intonation and the linguistic information conveyed in a discourse... without changing the nature of the speech people typically produce in conversation" (p.233). See the references above for examples of such tasks.

4 The perception and interpretation of prosodically realized focus

This section examines how listeners perceive and interpret prosodically realized focus.[4] Section 4.1 examines the perception and interpretation of prosodically realized focus across languages, and 4.2 shows that the perception and interpretation of prosodically realized focus is context-dependent. Section 4.3 illustrates that information-structural differences between foci are perceived and interpreted, and section 4.4 addresses the perception and interpretation of focus size. Section 4.5 concludes with a discussion of methods.

4.1 The perception and interpretation of focus across languages

Not only do speakers use prosody to signal what is focused in an utterance (section 3.1), listeners also attend to prosodic cues to focus in their interpretation of utterances. In an early study on American English, Most and Saltz (1979) found that listeners rely on the location of pitch accents in utterances of sentences

[4]The research reviewed in this section is limited to experimental research, to the exclusion of works that analyze the meaning of (mostly English) prosody by abstracting over findings in the experimental literature (e.g., Pierrehumbert and Hirschberg 1990, Steedman 2000, 2004, 2014 and Truckenbrodt 2012). Specifically, the research reviewed in this section is limited to works based on perception tasks, which ask participants to make a judgment about the signal (e.g., prominence ratings), and interpretation tasks, which ask participants to make a judgment about meaning (e.g., whether an utterance is congruent with a question).

like *The pitcher threw the ball* in identifying whether the subject or the object is focused. In one of their experiments, participants listened to utterances of such sentences in which either the subject or the object was prosodically prominent and were asked to write a question to which the utterance would be an appropriate answer. Most and Saltz found that the expression that was prominent in the answer utterance corresponded, in 68% of cases, to the expression that was focused in the utterance, given the question that the participant had provided. For instance, the question *Who threw the ball?* was written for the utterance *The pitcher threw the ball* in which the subject was prominent. Birch and Clifton (1995) also relied on question-answer congruence to explore whether listeners attend to prosody in identifying foci. Asking native speakers of American English to judge the appropriateness of congruent and incongruent question-answer pairs, they found that listeners rated utterances in which a focused noun phrase was pitch accented as more appropriate than utterances in which a focused noun phrase was not pitch accented. Breen et al. (2010) also found that listeners in their experiment 2 (where speakers and listeners received feedback about whether the listeners' response was correct) were able to determine whether the subject, verb or object was focused. In sum, these studies provide empirical evidence that listeners attend to prosodic cues in identifying which expression of an utterance is focused.[5]

Clopper and Tonhauser (2013) investigated whether Paraguayan Guaraní listeners attended to prosodic cues to focus, as well as which prosodic cues they attended to. As mentioned in section 3.1, Clopper and Tonhauser identified, on the basis of production experiments, that focused expressions in Guaraní are more likely than non-focused expressions to be realized with a rising pitch accent, with a longer duration and a steeper slope for rising pitch accents. The question-answer pairs collected in one of the production experiments served as stimuli for an interpretation and a perception task. In the interpretation task, Guaraní listeners were asked to identify which of two lexically identical answers that differed in their prosody was the preferred response to a question. They found that listeners attended to the pitch accent type and the slope of rising pitch accents in identifying which expression in the utterance was focused. However, they also found that Guaraní listeners' performance was significantly lower than chance on this task, i.e., listeners were not able to reliably use prosodic cues to identify which expression in an utterance was focused, contrary to what, e.g., Breen et al. 2010 found. As discussed in Clopper and Tonhauser 2013:§4.6, this result may at least partially be attributed to the task design (listeners did not receive feedback about whether their answers were correct) and to the fact that the materials included highly variable natural Guaraní utterances that included incongruous prosodic cues to focus.

In the prominence perception task, Clopper and Tonhauser asked Guaraní listeners to identify which word of the two-word answer utterances was prosodically more prominent. They found that listeners attended to the duration of the stressed syllable in identifying which expression was perceived as prosodically more prominent. As in the interpretation task, performance was not significantly above chance: listeners judged the focused expression to be more prosodically prominent than the non-focused one in 52% of cases. As discussed in Clopper and Tonhauser 2013:§5.5, the chance performance is consistent with research examining the perception of prosodic prominence by naïve listeners in other languages (e.g., Streefkerk et al. 1997, Mo et al. 2008). Taken together, the results from interpretation and perception tasks suggest that prosodic cues to focus constrain but do not determine what is focused in an utterance.

In sum, there is evidence from genetically unrelated languages that listeners attend to prosodic cues in identifying what is focused in an utterance. Perception and interpretation experiments can also serve to identify which prosodic cues to focus listeners attend to, as well as how succesful listeners are in determining from prosodic cues which expression is focused. Whether languages differ in the strength of their prosodic cues to focus is an open question.

[5] Accent placement facilitates listeners' utterance comprehension (e.g., Bock and Mazzella 1983, Terken and Nooteboom 1987, Ayers 1996) and information about focus from accent placement is processed rapidly online (e.g., Dahan et al. 2002). For a more thorough discussion of the influence of prosodic prominence on processing and memory see Kim's chapter in this volume.

4.2 Contextual influence on the perception and interpretation of prosodically realized focus

Context not only has an effect on the production of prosodically realized focus (section 3.2), but also on its perception and interpretation. Bishop (2012), for instance, showed that American English listeners' perception of prosodically realized focus is influenced by the question that the utterance addresses. In his study, participants listened to question-answer pairs, like the three shown in (7), and were asked to rate how "stressed" each word in the answer utterances sounded. Crucially, the same recording was used for the answer in each question-answer pair, namely a recording of the answer with verb phrase focus.

(7) A: What happened yesterday? / What did you do yesterday? / What did you buy yesterday?
B: I bought a motorcycle.

Bishop found that the objects of the answer utterances were rated as more stressed when the answers were heard in response to an object focus question than in response to a verb phrase or sentence focus question. Thus, "listeners' judgments of prosodic prominence are significantly and independently affected by their interpretation of an utterance's information structure" (p.255), i.e., by what is contextually identified as the focus of the utterance. Similarly, Turnbull et al. (2014) found that American English expressions that were realized with L+H* pitch accents were more likely to be perceived as prosodically prominent when listeners were explicitly made aware of the contrastive context in which the utterances were made.

The perception of prosodic prominence is also influenced by the prosodic context. In Krahmer and Swerts's (2001) study, Dutch listeners were asked to distinguish between words realized with (what the authors refer to as) contrastive and new accents. When presented with expressions realized with such accents in their sentential context, expressions with contrastive accents were rated as more prominent than expressions with new accents, but this difference in prominence rating was not observed when the expressions were presented in isolation. Thus, listeners attended to prosodic cues to the information-structural status of the expressions, but these cues were contextually enhanced.

Context also influences the interpretation of prosodically realized focus, as shown by Kurumada et al. (2012). Their study investigated the interpretation of prosodically realized focus using sentences of the form *It looks like a...*, such as *It looks like a zebra*. Utterances of these sentences differ in their interpretation depending on whether the verb *looks* or the noun (e.g., *zebra*) is the focus: while the utterances typically favor an affirmative interpretation ('it's a zebra') regardless of which expression is focused, the negative interpretation ('it's not a zebra') is more likely to arise when the verb is focused than when the noun is focused. In one experiment, Kurumada et al. (2012) showed that when listeners were made aware that speakers can also produce *It's a zebra*, they were more likely to assign negative interpretations to *It looks like a zebra*, regardless of focus placement. This finding suggests that contextual information about the speaker's intended meaning can weaken the contribution of prosodically realized focus to meaning. Another experiment reported in Kurumada et al. 2012 suggests that information about a speaker's use of prosody to convey meaning influences the extent to which listeners rely on prosodically realized focus in deriving the speaker's intended meaning.

In conclusion, the studies reviewed above show that the perception and interpretation of focus is context-dependent — at least in English and other well-studied European languages.

4.3 Interpreting prosodic cues to foci that differ in other information-structural properties

The studies reviewed in this section show that listeners attend not just to prosodic cues to focus (section 4.1) but also to prosodic cues to other information-structural properties of foci. Breen et al. (2010), for instance, found that a few American English listeners were able to rely on prosodic cues to distinguish whether an utterance like *Damon fried an omelet this morning* was made in response to a question like *Who fried an omelet this morning?* (in which *Damon* is new information) or to a question like *Did Harry fry an omelet*

this morning? (in which *Damon* is corrective and contrasted with *Harry*). They found that more listeners were able to identify which question the utterance was responding to when the utterance was preceded by *I heard that...*, which suggests that "speakers are able to convey contrastiveness using words outside of the clause containing the contrastively focused element" (p.1089).

Many studies have compared the interpretation of high tone pitch accents (H* in the ToBI annotation system, see footnote 3) and L+H* accents. In one of Ito and Speer's (2008) eye tracking experiments, listeners heard sequences like *... green drum... blue drum* and were found to fixate on the relevant objects (e.g., drums) earlier when *blue* was realized with a L+H* accent than with a H* accent. This finding suggests that listeners are more likely to interpret adjectives as contrastively focused when they are realized with a L+H* accent than when they are realized with a H* pitch accent. Likewise, Gotzner and Spalek (2014) showed, using a truth value judgment task, that German expressions were more likely to generate exhaustive inferences when realized with L+H* rather than H* accents. The question of whether H* marks foci whose denotation is new (as claimed in, e.g., Pierrehumbert and Hirschberg 1990), was addressed in Watson et al.'s (2008) eye tracking study. While they did find that American English listeners were more likely to fixate on a contrastive referent than on a new one when they heard the L+H* accent, listeners were not more likely to fixate on a new referent than a contrastive one when they heard the H* accent (an expression was taken to be 'contrastive' if it contrasted with the denotation of an expression in an immediately prior utterance, similarly to Ito and Speer 2008). Watson et al. (2008) took these findings to show that L+H* conveys a contrastive meaning (in support of Pierrehumbert and Hirschberg 1990) but that H* is compatible with new and contrastive foci.

Although the aforementioned studies suggest that listeners interpret H* and L+H* differently, other research suggests that the two pitch accents do not correspond to two distinct perceptual categories. Bartels and Kingston (1994), for instance, synthesized a continuum of stimuli between H* and L+H* by varying the acoustic characteristics of the accent (peak height, peak alignment, peak timing, depth of the dip preceding the peak) and the context in which the sentences were presented. They asked American English listeners to rate whether the relevant expression received a contrastive or a non-contrastive interpretation. (Here, contrastiveness was defined as the explicit denial of a salient alternative, like *apple* in *Amanda had a banana. – No, she had an apple*.) Bartels and Kingston did not find clear evidence for a categorical boundary between L+H* and H*. What they did find was that the higher the peak height was, the more likely a contrastive interpretation was. (But see Calhoun 2004 for a different finding.) The depth of the dip preceding the peak and the timing of the peak were secondary cues to a contrastive interpretation. Ladd and Morton (1997) found that listeners categorically distinguished the two accents in an identification task but they did not identify a category boundary in discrimination tasks, which they took to mean that the two accents are categorically interpreted, but continuously perceived.

4.4 Perceiving and interpreting the size of prosodically realized focus

Listeners can perceive and interpret prosodic cues to differences in focus size. D'Imperio (1997), for instance, found, using a prominence perception task, that Neapolitan Italian listeners perceive utterances of subject-verb-object declarative sentences with sentence focus differently from utterances with object focus: listeners identified the object as most prominent in only 56% of sentence focus utterances but in 77% of object focus utterances.

Can listeners interpret prosodic cues to differences in focus size? Gussenhoven (1983) answered this question negatively. He found that participants were at chance in distinguishing whether an utterance of a subject-verb-object sentence, like *We repair radios*, that had been made in response to a question that elicited object focus (*What is it you repair?*) responded to that question or to a question about the verb phrase (*What is the nature of your business?*). Breen et al. (2010), on the other hand, found that American English listeners can distinguish between object and sentence focus: listeners in their experiment 2 were able to correctly

identify utterances that were produced with (non-contrastive) object focus 57% of the time, perceiving such utterances as having sentence focus only 13% of the time. Similarly to Gussenhoven's (1983) task, listeners in their study had to identify which question an utterance responded to, but in contrast to Gussenhoven's task, listeners in their study were seated in the same room as the talkers, and both the talkers and the listeners received feedback when a listener chose the wrong question: "This cue was included to ensure that speakers knew when their productions did not contain enough information for the listener to choose the correct answer" (Breen et al. 2010:1067). Since the talkers' utterances more clearly disambiguated focus size when such feedback was given, it is possible that the listeners' ability to perceive distinctions in focus size depends on whether talkers are aware of the need to disambiguate their productions.

As noted in Bishop 2017, which provides an overview of research into the perception and interpretation of prosodic cues to focus size, previous studies seem to suggest that prenuclear prominence (i.e., a pitch accent on the verb) may be "less relevant to identifying broad focus than to identifying narrow focus" (p.238). The results of his investigation into the processing of prosodic cues to focus size, which involved a cross-modal priming task, suggest that prenuclear accents were unexpected when the object of subject-verb-object sentences was narrowly focused, but did not influence reaction times when the verb phrase of the sentence was focused. While Bishop's findings are compatible with Selkirk's (1996) Focus Projection hypothesis, studies like D'Imperio 1997 or Breen et al. 2010 challenge the assumption that an utterance with a prosodically prominent object noun phrase is ambiguous between object, verb phrase and sentence focus.

4.5 Experimental tasks for the perception and interpretation of prosodically realized focus

Although researchers have intuitions about how they perceive and interpret prosodically realized focus, the preceding sections have shown that experimental investigations are essential to identifying the phonetic and phonological properties of utterances that naïve listeners attend to. Experimental investigations can also identify inter-speaker variation in the prosodic cues attended to; for discussion see, e.g., Breen et al. 2010 and Bishop 2017. The preceding sections have also highlighted that research at the interface of prosody and meaning must carefully define the meanings that are investigated, including 'focus' or 'contrastive focus'. Finally, research must consider that the perception and interpretation of prosodically realized meaning, such as focus, may be context-dependent and control for context: when context is not controlled for, participants' interpretations may differ simply because participants imagine different contexts.

Experimental research on the perception and interpretation of prosodically realized focus employs a wide range of tasks: participants are asked to identify important or prosodically prominent words (e.g., D'Imperio 1997, Clopper and Tonhauser 2013), to identify congruent question-answer pairs (e.g., Most and Saltz 1979, Gussenhoven 1983, Birch and Clifton 1995, Breen et al. 2010, Clopper and Tonhauser 2013), or to answer questions about what a particular utterance means (e.g., Kurumada et al. 2012), or whether the utterance is true (e.g., Terken and Nooteboom 1987). Eye tracking and priming experiments provide information about the processing of prosodically realized focus (e.g., Dahan et al. 2002, Ito and Speer 2006, Bishop 2017). Stimuli can be resynthesized utterances (e.g., Ladd and Morton 1997, Calhoun 2004) or utterances produced by trained speakers (e.g., Gussenhoven 1983) or untrained speakers (e.g., Breen et al. 2010, Clopper and Tonhauser 2013); they may also differ in whether they realize specific prosodies (e.g., D'Imperio 1997, Ito and Speer 2006) or are randomly selected from a production experiment (e.g., Breen et al. 2010, Clopper and Tonhauser 2013). See Watson et al. (2006) for arguments in favor of experimental tasks that can be used to explore both the production and comprehension of prosody, and tasks that allow for the integration of contextual information.

5 Concluding remarks

In this article, an information-structural focus was defined as an expression that answers the question addressed by the utterance in which the expression is realized. As discussed in the preceding sections, the prosodic realization of utterances in many languages constrain what is focused in the utterance, and speakers of these languages can perceive and interpret prosodically realized foci. Experimental investigations are invaluable in identifying the phonetic and phonological cues to focus because such cues are subject to inter-speaker and inter-utterance variation, and because the phonetic and phonological properties of utterances are not only implicated in conveying focus. Although many open questions remain, the results of this research already have important implications for theories of meaning. For instance, theories of meaning must consider that the prosodic realization of focus is highly variable across utterances, speakers and contexts. Similarly, although the prosodic realization of an utterance provides cues to focus, prosodically realized focus does not, typically, unambiguously determine the perception and interpretation of focus in such utterances, even when the prosodic cues to focus are consistent and speakers were aware of the ambiguity to be resolved. As a consequence, theories of meaning that start from the premise that the input to interpretation unambiguously marks what is focused in an utterance (e.g., via F-marking in the syntax) constitute an idealization of the complex interplay of prosody and focus. (Discrepancies between empirical findings and theoretical analyses in research on clefts are discussed in Onea's chapter in this volume.)

Information-structural focus is related to or implicated in other categories of meaning. As a consequence, studies of the production, perception and interpretation of focus are relevant for a broad set of meanings beyond focus. Contrastive topics, for instance, have been defined as expressions that answer super-questions to the question addressed by the utterance in which the expression is realized, i.e., as foci relative to a super-question (e.g., Jackendoff 1972, Roberts 2012/1998, Büring 2003). It is therefore not surprising that contrastive topics, too, can be prosodically realized, and that prosodically realized contrastive topics can be perceived and interpreted (e.g., Braun 2006, Wagner et al. 2013). Contrastive topic-marking in turn influences the interpretation of sentences with scalar adjectives: de Marneffe and Tonhauser (to appear) found that American English utterances with scalar adjectives that were realized with the contrastive topic contour were more likely to give rise to a scalar implicature than when the utterance was realized with (what they referred to as) a "neutral" contour. More generally, research on conversational implicatures has established, for a range of scalar expressions, that the relevant scalar implicature is more likely to arise when the scalar expression is prosodically prominent than when it is not (e.g., Chevallier et al. 2008, Thorward 2009, Zondervan 2010, Cummins and Rohde 2015, Schwarz et al. to appear). Lastly, it has been hypothesized that information-structural focus is implicated in presupposition projection (e.g., Beaver 2010, Beaver et al. 2017, Simons et al. 2017), and experimental research on presuppositions has shown that the prosodic realizations of utterances with presupposition triggers influences the projectivity of the presuppositions (Cummins and Rohde 2015, Tonhauser 2016, Stevens et al. 2017); see also the discussion in Schwarz's chapter on presuppositions in this volume. In conclusion, as these few examples already show, experimental investigations of prosody are essential to developing a comprehensive perspective on natural language meaning.

Acknowledgments

I thank Cynthia Clopper, Chris Cummins, Craige Roberts and Jon Stevens for comments on earlier drafts of this article, and I gratefully acknowledge support from *National Science Foundation* grant BCS-1452674 in the writing of this article.

References

Ayers, Gayle M. 1996. *Nuclear accent types and prominence: Some psycholinguistic experiments*. Ph.D. thesis, The Ohio State University.

Bartels, Christine and John Kingston. 1994. Salient pitch cues in the perception of contrastive focus. In P. Bosch and R. van der Sandt, eds., *Focus and Natural Language Processing (IBM Working Papers on Logic and Linguistics, v. 6)*, pages 1–10. Columbia: University of Missouri.

Baumann, Stefan, Martine Grice, and Susanne Steindamm. 2006. Prosodic marking of focus domains: Categorical or gradient? In *Proceedings of Speech Prosody 2006*, pages 301–304. Dresden, Germany.

Beaver, David. 2010. Have you noticed that your belly button lint colour is related to the colour of your clothing? In R. Bäuerle, U. Reyle, and E. Zimmermann, eds., *Presuppositions and Discourse: Essays Offered to Hans Kamp*, pages 65–99. Oxford: Elsevier.

Beaver, David and Brady Clark. 2008. *Sense and Sensitivity: How Focus Determines Meaning*. Oxford: Wiley-Blackwell.

Beaver, David, Brady Clark, Edward Flemming, Florian Jaeger, and Maria Wolters. 2007. When semantics meets phonetics: Acoustical studies of second occurrence focus. *Language* 83:245–276.

Beaver, David, Craige Roberts, Mandy Simons, and Judith Tonhauser. 2017. Questions Under Discussion: Where information structure meets projective content. *Annual Review of Linguistics* 3:265–284.

Beckman, Mary E. and Jennifer J. Venditti. 2011. Intonation. In J. Goldsmith, J. Riggle, and A. Yu, eds., *Handbook of Phonological Theory*, pages 485–532. Malden MA: Blackwell.

Beckman, Mary E. 1996. The parsing of prosody. *Language and Cognitive Processes* 11:17–67.

Beckman, Mary E. and Gayle M. Ayers. 1997. *Guidelines for ToBI labelling, version 3.0*. The Ohio State University.

Birch, Stacy and Charles Clifton. 1995. Focus, accent, and argument structure: Effects on language comprehension. *Language and Speech* 38:365–391.

Bishop, Jason. 2012. Information structural expectations in the perception of prosodic prominence. In G. Elordieta and P. Prieto, eds., *Prosody and Meaning*, pages 239–269. Berlin: Mouton de Gruyter.

Bishop, Jason. 2017. Focus projection and prenuclear accents: evidence from lexical processing. *Language, Cognition and Neuroscience* 32:236–253.

Bock, Kathryn J. and Joanne R. Mazzella. 1983. Intonational marking of given and new information: Some consequences for comprehension. *Memory & Language* 11:64–76.

Braun, Bettina. 2006. Phonetics and phonology of thematic contrast in German. *Language and Speech* 49:451–493.

Breen, Mara, Evelina Fedorenko, Michael Wagner, and Edward Gibson. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes* 25:1044–1098.

Brosch, Tobias, Gilles Pourtois, and David Sander. 2010. The perception and categorisation of emotional stimuli: A review. *Cognition and Emotion* 24:377–400.

Burdin, Rachel Steindel, Sara Phillips-Bourass, Rory Turnbull, Murat Yasavul, Cynthia G. Clopper, and Judith Tonhauser. 2015. Variation in the prosody of focus in head- and head/edge-prominence languages. *Lingua* 165:254–276.

Büring, Daniel. 2003. On D-trees, beans and B-accents. *Linguistics & Philosophy* 26:511–545.

Büring, Daniel. 2007. Intonation, semantics and information structure. In G. Ramchand and C. Reiss, eds., *The Oxford Handbook of Linguistic Interfaces*, pages 445–473. Oxford: Oxford University Press.

Calhoun, Sasha. 2004. Phonetic dimensions of intonational categories: The case of L+H* and H*. In *Proceedings of the International Conference on Spoken Language Processing*. Nara: Japan.

Calhoun, Sasha. 2010. The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language* 86:1–42.

Chevallier, Coralie, Ira A. Noveck, Tatjana Nazir, Lewis Bott, Valentina Lanzetti, and Dan Sperber. 2008.

Making disjunctions exclusive. *The Quarterly Journal of Experimental Psychology* 61:1741–1760.

Clopper, Cynthia G. and Judith Tonhauser. 2011. On the prosodic coding of focus in Paraguayan Guaraní. In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, pages 249–257. Somerville, MA: Cascadilla Proceedings Project.

Clopper, Cynthia G. and Judith Tonhauser. 2013. The prosody of focus in Paraguayan Guaraní. *International Journal of American Linguistics* 79:219–251.

Cooper, William E., Stephen Eady, and Pamela R. Mueller. 1985. Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America* 77:2142–2156.

Cummins, Chris and Hannah Rohde. 2015. Evoking context with contrastive stress: Effects on pragmatic enrichment. *Frontiers in Psychology* 6:Article 1779.

Dahan, Delphine, Michael K. Tanenhaus, and Craig G. Chambers. 2002. Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language* 47:292–314.

de Marneffe, Marie-Catherine and Judith Tonhauser. to appear. Inferring meaning from indirect answers to polar questions: The contribution of the rise-fall-rise contour. In E. Onea, M. Zimmermann, and K. von Heusinger, eds., *Questions in Discourse*. Leiden: Brill.

D'Imperio, Mariapaola. 1997. Breadth of focus, modality and prominence perception in Neapolitan Italian. *Ohio State University Working Papers in Linguistics* 50:19–39.

D'Imperio, Mariapaola. 2001. Focus and tonal structure in Neapolitan Italian. *Speech Communication* 33:339–356.

Downing, Laura. 2008. Focus and prominence in Chichewa, Chitumbuka and Durban Zulu. *ZAS Papers in Linguistics (Papers in Phonetics and Phonology)* 49:47–65.

Eady, Stephen and William E. Cooper. 1986. Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America* 80:402–415.

Eady, Stephen J., William E. Cooper, Gayle V. Klouda, Pamela R. Mueller, and Dan W. Lotts. 1986. Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments. *Language and Speech* 29:233–251.

Féry, Caroline. 2007. Information structural notions and the fallacy of invariant correlates. In C. Féry, G. Fanselow, and M. Krifka, eds., *Interdisciplinary Studies on Information Structure 6*, pages 161–184. Potsdam.

Féry, Caroline. 2013. Focus as prosodic alignment. *Natural Language & Linguistic Theory* 31:683–734.

Féry, Caroline and Manfred Krifka. 2008. Information structure: Notional distinctions, ways of expression. In P. van Sterkenburg, ed., *Unity and Diversity of Languages*, pages 123–136. Amsterdam: John Benjamins.

Frota, Sónia. 2002. The prosody of focus: A case-study with cross-linguistic implications. In *Speech Prosody 2002*, pages 319–322. Aix-en-Provence, France.

Gotzner, Nicole and Katharina Spalek. 2014. Exhaustive inferences and additive presuppositions: Interplay of focus operators and contrastive intonation. In *Proceedings of the Formal & Experimental Pragmatics Workshop*, pages 7–13. Tübingen.

Grice, Martine and Stefan Baumann. 2008. An introduction to intonation – functions and models. In J. Trouvain and U. Gut, eds., *Non-Native Prosody. Phonetic Description and Teaching Practice*, pages 25–51. Berlin, New York: De Gruyter.

Gunlogson, Christine. 2002. Declarative questions. In *Proceedings of Semantics and Linguistic Theory (SALT) XII*, pages 124–143. Ithaca, NY: CLC Publications.

Gussenhoven, Carlos. 1983. Testing the reality of focus domains. *Language and Speech* 26:61–80.

Gussenhoven, Carlos. 2002. Phonology of intonation. *Glot International* 6:271–284.

Halliday, Michael A. K. 1985. *An Introduction to Functional Grammar*. London and Baltimore, Md.: Edward Arnold.

Hanssen, Judith, Jörg Peters, and Carlos Gussenhoven. 2008. Prosodic effects of focus in Dutch declaratives.

In *Proceedings of Speech Prosody 2008*, pages 609–612. Campiñas, Brazil.

Hirst, Daniel and Albert Di Cristo. 1998. *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press.

Ito, Kiwako and Shari R. Speer. 2006. Using interactive tasks to elicit natural dialogue. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, and J. Schließer, eds., *Methods in Empirical Prosody Research*, pages 229–257. New York: Walter de Gruyter.

Ito, Kiwako, Shari R. Speer, and Mary E. Beckman. 2004. Informational status and pitch accent distribution in spontaneous dialogues in English. In *Proceedings of the International Conference on Spoken Language Processing*, pages 279–282. Nara: Japan.

Ito, Kiwako and Shari R. Speer. 2008. Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language* 58:541–573.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jacobs, Joachim. 1983. *Fokus und Skalen*. Tübingen: Niemeyer.

Jeong, Sunwoo and Christopher Potts. 2016. Intonational sentence-type conventions for perlocutionary effects: An experimental investigation. In *Semantics and Linguistic Theory (SALT) 26*, pages 1–22. eLanguage.

Jun, Sun-Ah. 1993. *The Phonetics and Phonology of Korean Prosody*. Ph.D. thesis, The Ohio State University.

Jun, Sun-Ah. 2000. K-ToBI (Korean ToBI) labelling conventions (ver. 3.1). *UCLA Working Papers in Phonetics* 99:149–173.

Jun, Sun-Ah, ed. 2005. *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Jun, Sun-Ah. 2014. Prosodic typology: By prominence type, word prosody, and macro-rhythm. In S.-A. Jun, ed., *Prosodic Typology*. Oxford: Oxford University Press.

Jun, Sun-Ah and Hee-Sun Kim. 2008. VP focus and narrow focus in Korean. In *Proceedings of the XVIth International Congress of Phonetic Sciences*, pages 1277–1280. Saarbrücken, Germany.

Jun, Sun-Ah and Hyuck-Joon Lee. 1998. Phonetic and phonological markers of contrastive focus in Korean. In *Proceedings of the 5th International Conference on Spoken Language Processing*, pages 1295–1298. Sydney, Australia.

Katz, Jonah and Elisabeth Selkirk. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language* 87:771–816.

Kim, Hee-Sun, Sun-Ah Jun, Hyuck-Joon Lee, and Jong-Bok Kim. 2006. Argument structure and focus projection in Korean. In *Proceedings of Speech Prosody 2006*. Dresden, Germany.

Kim, Sahyang. 2004. *The Role of Prosodic Phrasing in Korean Word Segmentation*. Ph.D. thesis, UCLA.

Kiss, Katalin É. 1998. Identificational focus versus information focus. *Language* 74:245–273.

Koch, Karsten. 2008. *Intonation and Focus in Nɬeʔkepmxcin*. Ph.D. thesis, University of British Columbia.

Krahmer, Emiel and Marc Swerts. 2001. On the alleged existence of contrastive accents. *Speech Communication* 34:391–405.

Krifka, Manfred. 2006. Basic notions of information structure. In C. Féry, G. Fanselow, and M. Krifka, eds., *Interdisciplinary Studies of Information Structure 6*, pages 13–55. Potsdam: Universitätsverlag Potsdam.

Kügler, Frank and Stavros Skopeteas. 2006. Interaction of lexical tone and information structure of Yucatec Maya. In *Second International Symposium on Tonal Aspects of Languages*, pages 380–388. La Rochelle, France.

Kügler, Frank and Stavros Skopeteas. 2007. On the universality of prosodic reflexes of contrast: The case of Yucatec Maya. In *16th International Congress of the Phonetic Sciences (ICPhS)*, pages 1025–1028.

Kurumada, Chigusa, Meredith Brown, and Michael K. Tannenhaus. 2012. Pragmatic interpretation of contrastive prosody: It looks like speech adaptation. In *Proceedings of the 35th Annual Conference of*

the Cognitive Science Society. Sapporo, Japan: Cognitive Science Society.

Ladd, D. Robert. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press, 2nd edn.

Ladd, Robert D. and Rachel Morton. 1997. The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics* 25:313–342.

Lambrecht, Knut. 1994. *Information Structure and Sentence Form*. Cambridge, England: Cambridge University Press.

Mo, Yoonsook, Jennifer Cole, and Eun-Kyung Lee. 2008. Naïve listeners' prominence and boundary perception. In *Speech Prosody 2008*, pages 735–38. Campinas.

Most, Robert B. and Eli Saltz. 1979. Information structure in sentences: New information. *Language and Speech* 22:89–95.

Paul, Hermann. 1880. *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer.

Paul, Hermann. 1919. *Deutsche Grammatik*, vol. Volume 3. Tübingen: Niemeyer.

Pierrehumbert, Janet and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In J. M. P. Cohen and M. Pollack, eds., *Intentions in Communication*, pages 271–311.

Prieto, Pilar. 2015. Intonational meaning. *Wiley Interdisciplinary Reviews: Cognitive Science* 6:371–381.

Rialland, Annie and Stéphane Robert. 2001. The intonational system of Wolof. *Linguistics* 39:893–939.

Roberts, Craige. 1998. Focus, information flow, and Universal Grammar. In P. Culicover and L. McNally, eds., *Syntax and Semantics Volume 29: The Limits of Syntax*, pages 109–160. New York: Academic Press.

Roberts, Craige. 2012. Topics. In C. Maienborn, P. Portner, and K. von Heusinger, eds., *Handbook of Semantics*, pages 1908–1934. Berlin: Mouton de Gruyter.

Roberts, Craige. 2012/1998. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 5:1–69. Reprint of 1998 revision of 1996 OSU Working Papers publication.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.

Schwarz, Florian, Chuck Clifton, and Lyn Frazier. to appear. Strengthening 'or': Effects of focus and downward entailing contexts on scalar implicatures. In F. S. J. Anderssen, K. Moulton and C. Ussery, eds., *Semantics and Processing (UMOP 37)*. Amherst, MA: GLSA Publications.

Schwarzschild, Roger. 1999. Givenness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7:141–177.

Selkirk, Elizabeth. 2007. Contrastive focus, givenness and the unmarked status of "discourse-new". *Acta Linguistica Hungarica* 55:331–346.

Selkirk, Elizabeth O. 1996. Sentence prosody: Intonation, stress, and phrasing. In J. Goldsmith, ed., *The Handbook of Phonological Theory*, pages 550–569. Blackwell.

Shattuck-Hufnagel, Stephanie and Alice E. Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25:193–246.

Simons, Mandy, David Beaver, Craige Roberts, and Judith Tonhauser. 2017. The Best Question: Explaining the projection behavior of factive verbs. *Discourse Processes* 3:187–206.

Speer, Shari R., Paul Warren, and Amy J. Schafer. 2011. Situationally independent prosodic phrasing. *Laboratory Phonology* 2:35–98.

Steedman, Mark. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry* 31:649–689.

Steedman, Mark. 2004. Information-structural semantics for English intonation. In C. Lee, M. Gordon, and D. Büring, eds., *Topic and Focus: Crosslinguistic Perspectives on Meaning and Intonation*, pages 247–266. Dordrecht: Springer.

Steedman, Mark. 2014. The surface-compositional semantics of English intonation. *Language* 90:1–57.

Stevens, Jon Scott, Marie-Catherine de Marneffe, Shari R. Speer, and Judith Tonhauser. 2017. Rational use of prosody predicts projectivity in manner adverb utterances. In *39th Annual Meeting of the*

Cognitive Science Society.

Streefkerk, Barbertje, Louis C W Pols, and Louis F M ten Bosch. 1997. Prominence in read aloud sentences, as marked by listeners and classifiers automatically. In *Proceedings*, vol. 21, pages 101–16. Institute of Phonetic Sciences, University of Amsterdam.

Terken, Jacques and Julia Hirschberg. 1994. Deaccentuation of words representing 'given' information: Effects of persistence of grammatical function and surface position. *Language and Speech* 37:125–145.

Terken, Jacques and Sieb G. Nooteboom. 1987. Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes* 2:145–163.

Thorward, Jennifer. 2009. The interaction of contrastive stress and grammatical context in child English speakers' interpretations of existential quantifiers. B.A. thesis, The Ohio State University.

Tonhauser, Judith. 2016. Prosodic cues to speaker commitment. In *Semantics and Linguistic Theory (SALT) XXVI*, pages 934–960. Ithaca, NY: CLC Publications.

Truckenbrodt, Hubert. 2012. Semantics of intonation. In C. Maienborn, K. von Heusinger, and P. Portner, eds., *Semantics: An International Handbook of Natural Language Meaning*, pages 2039–2969. Berlin: Mouton de Gruyter.

Turnbull, Rory, Rachel Steindel Burdin, Cynthia G. Clopper, and Judith Tonhauser. 2015. Contextual information and the prosodic realization of focus: A cross-linguistic comparison. *Language, Cognition and Neuroscience* 30:1061–1076.

Turnbull, Rory, Adam J Royer, Kiwako Ito, and Shari R Speer. 2014. Prominence perception in and out of context. In *Proceedings of the 7th International Conference on Speech Prosody*, pages 1164–1168.

Vallduvi, Enrique and Maria Vilkuna. 1998. On rheme and kontrast. In P. Culicover and L. McNally, eds., *The Limits of Syntax*, pages 79–108. New York: Academic Press.

von Stechow, Arnim. 1990. Focusing and backgrounding operators. In W. Abraham, ed., *Discourse Particles*, pages 37–84. Amsterdam: John Benjamins.

Wagner, Michael, Elise McClay, and Lauren Mak. 2013. Incomplete answers and the rise-fall-rise contour. In *Proceedings of the 17th Workshop on the Semantics and Pragmatics of Dialogue*, pages 140–149.

Wagner, Michael and Duane G. Watson. 2010. Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes* 25:905–945.

Watson, Duane G., Jennifer Arnold, and Michael K. Tanenhaus. 2008. Tic tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition* 106:1548–1557.

Watson, Duane G. 2010. The many roads to prominence: Understanding emphasis in conversation. *Psychology of Learning and Motivation* 52:163–183.

Watson, Duane G., Christine A. Gunlogson, and Michael K. Tanenhaus. 2006. Online methods for the investigation of prosody. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, and J. Schlieśer, eds., *Methods in Empirical Prosody Research*, pages 259–282. New York: Walter de Gruyter.

Xu, Yi and Ching X. Xu. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33:159–197.

Yasavul, Murat. 2013. Prosody of focus and contrastive topic in K'iche'. *Working Papers in Linguistics, The Ohio State University* 60:129–160.

Zerbian, Sabine. 2006. *Expression of Information Structure in the Bantu Language Northern Sotho*. Ph.D. thesis, Humboldt University, Berlin.

Zimmermann, Malte and Edgar Onea. 2011. Focus marking and focus interpretation. *Lingua* 121:1651–1670.

Zondervan, Albert Jan. 2010. *Scalar implicatures or focus: An experimental approach*. Ph.D. thesis, University of Utrecht.