31 Prosodic encoding of information structure: A typological perspective
Frank Kügler & Sasha Calhoun

31.1 Introduction
Information structure (IS) can be conceived of as a ‘cognitive domain’ that interacts with the linguistic modules syntax, phonology, and morphology on the one hand, and, on the other hand, with other cognitive capabilities that control, update and infer interlocutors’ common beliefs (Zimmermann & Féry 2010). A speaker’s utterance can be subdivided according to its IS, and the constituents of an utterance can be analysed as being focused, given, and/or topical. Interlocutors use different linguistic means to achieve the goal of controlling, updating and inferring their common beliefs. This chapter deals with the prosodic means that speakers use to signal the IS of an utterance from a typological perspective.

At first sight, it may appear that languages vary widely as to which prosodic cues signal IS (cf. Jun 2005b, 2014; Kügler 2011; Downing & Rialland 2016a). However, within prosody and IS research, it has recurrently been proposed that this variation in the prosodic realization of IS can be subsumed by underlying principles which point to common phonological structures present cross-linguistically (Truckenbrodt 1995; Büring 2010; Féry 2013). Another line of research concerns the debate whether there is a relationship between the prosodic profile of a language and its prosodic means to express IS (cf. Burdin et al. 2015). This chapter is organized around the different prosodic strategies to express IS found to date in the languages of the world (section 31.3), including interactions between syntax, prosody and IS (section 31.4); leading up to a discussion of commonalities between languages (section 31.5). Section 31.2 introduces the basic concepts of IS, and section 31.6 presents an evaluative outlook for future research.

31.2 Basic concepts of information structure
We adhere to the widely held view that the basic notions of IS are the cognitive categories focus, givenness and topic that are rooted in a theory of communication (Krifka 2008). IS refers to the division of an utterance in information packages (Chafe 1976) with the aim of a continuous information update of the common ground of interlocutors. The category focus is defined as an indication of ‘the presence of alternatives that are relevant for the interpretation of linguistic expressions’ (Krifka 2008: 247). This abstract cognitive category can be prosodically expressed in language-specific ways. Consider the mini-dialogue in (1). Speaker Q’s WH-question is a request to update the common ground with the referent of the WH-word. Speaker A responds by selecting the referent Peter as the relevant one of the possible alternatives in the context. This constituent is the focus and receives a pitch accent in languages like English, whereby the location of the pitch accent is given in upper case. The rest of the utterance is the background. The focused constituent in the answer corresponds to the WH-constituent in the question, thus creating a coherent discourse.

(1) Q: Who stole the cookie?  
A: [PEter]- stole the cookie.  (Krifka 2008: 250)

A further use of focus is to correct or confirm information that an interlocutor puts under discussion, as in (2). In this case, context (2C) contains a proposition containing potential focus alternatives. In (2A), the speaker corrects (2C), while (2A’) contains the identical proposition.
In (1), the focus is also new information in the context, and the background given. However, *givenness* is actually orthogonal to *focus* in that a focused constituent can either be *given* or *new*, which represent either ends of a givenness scale. *Givenness* refers to the information status of a constituent. A constituent is *given* if it is present in the immediate common ground (Krifka 2008: 262), or *new*, if it is not. A further, intermediate stage of the givenness scale is when the constituent is present in the common ground but may or may not be activated by the immediate discourse. In some languages, different degrees of saliency correlate with the degree of prosodic prominence a constituent is expressed with (see further section 31.3.1).

The example dialogue in (3) for instance shows a case where speaker A mentions the *cookie* and speaker B repeats the same word in the answer. The givenness status of *the cookie* affects accent assignment rules in some languages (Gussenhoven 1983; Selkirk 1995). The deaccentuation of *cookie* in combination with a focus accent on the verb constitutes discourse coherence and illustrates that a given constituent is expressed with less prosodic prominence.

Finally, the category *topic* refers to a constituent of an utterance that a speaker chooses to give further information about in the rest of the sentence, which is usually referred to as the *comment* (Krifka 2008: 264f). A topic in this sense is usually referred to as an *aboutness topic*. The division of a sentence into topic-comment opens a relationship between the information conveyed by the *topic* and the information given about the *topic*. Taking the answer of the cookie example (1) as a context-free sentence, (4) illustrates that *Peter* represents the *topic*, and the following information in the *comment* is about the topic constituent *Peter*.

The key IS categories are illustrated here in English, which relies heavily on prosodic cues to express IS. Note however that across languages linguistic devices other than prosodic means are used to express these information structural categories (Zimmermann & Féry 2010; Féry & Ishihara 2016). By far the prosodic expression of *focus* is most studied, but we discuss the expression of *givenness* and *topic* when possible. In many studies, the category of *focus/background* is conflated with referential status of *new/given*. We comment on work that separates these when relevant.

### 31.3 A typology of prosodic encoding of information structure

#### 31.3.1 Stress or pitch accent-based cues

Stress-based systems are the most well-studied type of prosodic encoding of IS, with decades of research, predominately on English. The basic pattern is that the focused word in an utterance is the most prosodically prominent (e.g. Selkirk 1995; Ladd 2008; Büring 2010, 2016; Calhoun 2010). Stress-based prominence is marked by phonetic and phonological cues which increase the prominence of a word relative to others in the utterance. Phonetic cues include higher fundamental frequency (f0), greater f0 movement, lengthening, increased intensity and higher spectral tilt on the word, as well as a drop in f0 after it (see summaries in Ladd 2008; Breen et al. 2010; Fletcher 2010; Turk 2011). Of these, f0 cues are the best studied and are perceptually important (see review in Cole 2015, and references therein).
Intensity and lengthening are also perceptually important, at least in English (Turk & Sawusch 1996; Kochanski et al. 2005; Breen et al. 2010).

<insert Figure 31.1 here>

Figure 31.1: Typical realisations of (1) and (4), showing how focus position affects prosodic realisation. A schematic pitch realization is given, along with the prosodic phrasing, intonational tune, and text, where capitals indicate the pitch accented syllable. See text for further details.

Phonologically, a word is the most prominent in an utterance because its main stressed syllable is the head of the largest prosodic phrase that it is part of (usually the intonation phrase, i) (Ladd 2008). The head of the t-phrase carries a nuclear pitch accent. Therefore, as established by work in the Autosegmental-Metrical (AM) framework, focus is not marked directly by phonetic cues, but rather indirectly: these cues primarily mark (nuclear) pitch accents, which mark focus (Selkirk 1995; Ladd 2008; Büring 2010, 2016; Calhoun 2010, inter alia). This is opposed to a direct view, where focus is taken to be marked directly by phonetic prominence (Eady et al. 1986; Xu & Xu 2005; Breen et al. 2010). Figure 31.1 illustrates typical realisations of (1) and (4). Each utterance has a prosodic phrase structure (phonological phrases, φ, and t-phrases, shown), and an intonational tune (according to the AM-ToBI scheme, see chapter 4 this volume). The nuclear accent is, by default, the final pitch accent in the t-phrase, hence on the focus Peter in Figure 31.1a, and cookie in Figure 31.1b. This is supported by perception studies, which show that listeners expect the nuclear accent, and therefore the focus, will be final, and that a final nuclear accent is compatible with variable focus scope (Ayers 1996; Terken & Hermes 2000; Carlson et al. 2009; Bishop 2012, 2017).

Pre-nuclear pitch accents may or may not mark additional foci in the sentence (Féry & Samek-Lodovici 2006; Calhoun 2010), e.g. the H* accent on Peter in Figure 31.1b does not. The prominence of pre-nuclear accents can also signal the distinction between broad and narrow focus; although there is substantial overlap between realisations compatible with each (Rump & Collier 1996; Ladd 2008: Ch. 7; Breen et al. 2010). Post-nuclear phonological heads are either not associated with pitch accents, i.e. they are ‘de-accented’, or if there are accents they appear in a compressed pitch register (Kügler & Féry 2017). Fully-fledged post-nuclear accents mark focus only in constrained discourse contexts such as second occurrence focus (see Beaver et al. 2007; Baumann 2016).

Stress-based marking of focus appears to be widespread amongst the world’s languages. It is ubiquitous amongst Germanic languages (Fanselow 2016), e.g. English, German (e.g. Féry & Kügler 2008), Dutch (e.g. Gussenhoven 2004) and Swedish (e.g. Bruce 1977; Myrberg & Riad 2016). Stress-based marking is also commonly found in Slavic languages (Jasinskaja 2016), including Russian (Luchkina & Cole 2016), and other Eastern European languages including Romanian (Manolescu et al. 2009), Greek (Baltazani & Jun 1999; Skopeteas 2016), Romani (Arvaniti & Adamou 2011) and Estonian (Ots 2017). Stress-based focus marking is also reported for Persian (Sadat-Tehrani 2007; Hosseini 2014), the Oceanic language Torau (Jepson 2014) and Paraguayan Guaraní (Clopper & Tonhauser 2013).

<insert Figure 31.2 here>
There are differences in the extent of stress-based marking of focus, however, even between closely related languages, or varieties of one language. For example, Chahal & Hellmuth (2014) report differences between Lebanese and Egyptian Arabic (see also Caron et al. 2015, and El Zarka 2017 for an overview of varieties of Arabic). In Lebanese Arabic, a narrowly focused word is nuclear accented, and any post-nuclear material deaccented (see Figure 31.2a); similar to English (see Figure 31.1). In Egyptian Arabic, every prosodic word (ω) is pitch accented, and there is no post-nuclear deaccentuation, although the pitch range of the focal accent is expanded (see Figure 31.2b). Romance languages are similar: some, including European Portuguese and Mexican, Argentinian, Basque and Canarian varieties of Spanish follow the standard stress-based pattern; whereas others including Madrid Spanish, Catalan and Italian only use in situ stress-based marking in some discourse contexts, e.g. contrastive focus, and post-nuclear accents are not deleted (Frota & Prieto 2015b; del Mar Vanrell & Fernández-Soriano 2018); see further sections 31.3.3 and 31.4. Along with stress-based marking, focus can usually also be signalled by word order, and this interacts with prosodic cues to focus (see section 31.4.). However, in the standard case, in situ stress-based marking, i.e. the nuclear accent can be in any position in the utterance, is an available, if not the preferred, means of marking focus (cf. Skopeteas & Fanselow 2010).

As discussed in section 2, referential givenness is orthogonal to focus; although focus and newness are correlated. In studies which separate focus and givenness, givenness is generally associated with lower stress-based prominence and/or deaccenting, both inside and out of focused constituents (e.g. Cutler et al. 1997; Baumann 2006; Féry & Kügler 2008; Baumann & Riester 2012; Cole et al. 2010; Baumann & Kügler 2015). As shown in (4), in English and other Germanic languages, within a multi-word focus, if the final word is highly contextually salient, it is usually deaccented (Wagner 2005; Ladd 2008; Riester & Piontek 2015). Similarly, given words in the background (non-focus) of an utterance are more likely to be unaccented than new words (Gussenhoven 1983; Selkirk 1995; Féry & Kügler 2008). Listeners typically find it harder to process discourses in which given items are accented, and new items unaccented (Cutler et al. 1997; Baumann & Schumacher 2012). Most of the empirical work in this area has been on Germanic languages, so it is unclear how widespread these patterns are. The tendency to deaccent given words within a focus, e.g. (4), is found for Slavic languages (Jasinskaja 2016) and Paraguayan Guarani (Burdin et al. 2015), but not in Romance languages (Ladd 2008; Swerts et al. 2002), or ‘outer circle’ varieties of English including Indian English and Caribbean English (Ladd 2008). Given words have lower prominence than new in unfocused positions in Hungarian (Genzel et al. 2015), but not Egyptian Arabic (Hellmuth 2011).

Types of nuclear accent or tone

Along with the placement of the nuclear accent, the pitch accent and/or boundary tone type is argued to play an important role in encoding IS in a variety of languages, particularly aspects of IS beyond focus, principally contrast and topic status.

One well-studied case is Romance languages, where there is widespread consensus that the nuclear accent + boundary tone configuration signals broad versus narrow/contrastive focus (see Frota & Prieto 2015b for overview); with the exception of French (see section 31.3.2.). A typical example is Sardinian (del Mar Vanrell et al. 2015). In broad focus, the nuclear accent
is usually H+L* (Figure 31.3a), while in narrow focus it is H*+L (Figure 31.3b), both
followed by a low boundary L%. That is, the nuclear accent peak is aligned earlier in broad
focus. In these languages, pre-nuclear accents are generally of one tonal type (L+H* in Figure
31.3), while nuclear accent (and following boundary tone) types vary and signal IS and other
pragmatic meanings.

Relatedly, in English and other Germanic languages, it is widely claimed that contrastive foci
are typically marked by L+H* accents, with H* for non-contrastive focus (e.g. Pierrehumbert
& Hirschberg 1990; Ito & Speer 2008; Watson et al. 2008). The L+H* and H* distinction,
has, however always been problematic, as it is difficult for annotators to make, and is the only
accent type distinction not based on the association of L and H tones. The most reliable cue to
the distinction is peak height (Ladd & Schepman 2003; Dilley 2005; Ladd 2008; Calhoun
2012; Repp 2016), which is experimentally confirmed for English (e.g. Welby 2003; Breen et
al. 2010; Katz & Selkirk 2011), Dutch (Krahmer & Swerts 2001) and German (Braun 2006;
Kügler & Gollrad 2015); suggesting this is better construed as a stress-based distinction.

Similar arguments have been made about Romance languages, i.e. the later peaks for narrow
focus are also often higher (e.g. see Figure 31.3), so peak alignment may be a partial proxy for
peak height (Gussenhoven 2004; Ladd 2008: Ch. 5; del Mar Vanrell et al. 2013; Borràs-
Comes et al. 2014; Repp 2016). It remains unsettled whether distinctions based on peak
scaling are gradient or categorical (e.g. Gussenhoven 2004; Ladd 2008: Ch. 5; Borràs-Comes
et al. 2014).

A connected issue is lack of consistency between studies as to what constitutes contrastive
focus: narrow focus (focus on a single word); contrastive (involving contextually identifiable
alternatives); or corrective (cf. (2)) (e.g. Repp 2016). It seems unlikely that any one language
systematically distinguishes all of these prosodically. Rather, it may be that speakers use
increased phonetic prominence to draw attention to foci that are contextually salient, e.g.
explicit contrast or corrections (Baumann et al. 2006; Calhoun 2009; Féry 2013).

In a number of languages, topics and foci are associated with different accent types: rising
accents (L+H*/L*+H) for topics and falling for foci (H*+L/L-), e.g. German (e.g. Braun
2006; Repp & Drenhaus 2015), English (Büring 2003; Steedman 2014), Russian (Alter &
Junghans 2002), and Arabic (El Zarka 2017). This is linked to sentence position, with topics
typically preceding foci. Calhoun (2012) claims for English the distinction is better construed
as one of relative prominence, with topics less relatively prominent than foci (see also Féry
2007).

Some proposals link a greater range of tonal event types with more detailed IS frameworks
(see also chapter 29, this volume). For example in English, Steedman (2000; 2014) proposes
four orthogonal dimensions of IS, signalled by combinations of pitch accent and boundary
tones: background/contrast (our focus), theme-rheme (roughly topic-comment), added/not
added to the common ground, speaker/hearer claimed (see also Brazil 1985; Pierrehumbert &
Hirschberg 1990; Gussenhoven 2004; Féry 1993 and Grice et al. 2005 for German; and Prieto
2015 for related work on Romance). It is unlikely, however, that there is a one-to-one
 correspondence between phonological tune types and IS expression (e.g. Féry 2008; Féry &
Ishihara 2016; Hirschberg et al. 2007; Zimmermann & Féry 2010).
The role of accent and boundary tone type in signalling IS has predominately been researched 
in languages with post-lexical pitch accents. However, they may have a role in other 
languages. For example, in Mandarin Chinese boundary tones on sentence final particles can 
signal meanings such as presupposition, which are part of IS (e.g. Peng et al. 2005, and 
references therein).

31.3.2 Phrase-based cues

The expression of focus by phrase-based cues first received more attention when research 
beyond well-studied languages like English contributed to the field. The basic insight is that 
the word-prosodic system of these languages predominantly does not have lexical prominence 
and hence focus cannot be expressed by pitch accenting as in languages with stress-based 
cues. The principal intonation units are phrase tones, and prosodic phrasing is the major 
component of intonation. The languages differ with respect to the domain that prosodic phrase 
markings indicate, e.g. a \( \omega \), a \( \varphi \)-phrase or an \( \iota \)-phrase, but the general commonality of these 
languages is that focus induces additional phrase boundaries. Hence, the highlighting function 
of focus is expressed by separating the focused constituent from other constituents; at the 
same time, post-focal constituents may be integrated into the same phrase as the focussed 
constituent.

Consider Korean which neither has lexical pitch accent nor lexical stress (cf. Jun 2005a). 
Korean distinguishes two levels of phrasing. The \( \iota \)-phrase dominates at least one accentual 
phrase (\( \alpha \)), and an \( \alpha \) consists of at least one \( \omega \). Each \( \alpha \) is tonally marked by two rising pitch 
patterns. The first one is associated with the initial two moras of the \( \alpha \) and the last one with 
the \( \alpha \)-final moras (5b). There is some variation as to the initial rise: if the \( \alpha \)-initial consonant 
is tense or aspirated the \( \alpha \)-initial tone is H (Jun 1998, 2005a). Phrase-final lengthening in 
combination with a low boundary tone demarcates the \( \iota \)-phrase, whereas the \( \alpha \) shows no final 
lengthening.

A prosodic boundary is consistently inserted before a focused constituent (Jeon & Nolan 
2017; Jun & Kim 2007; Jun & Lee 1998; Yang 2017), while following words tend to be in the 
same \( \varphi \) as the focused constituent (5c). Jeon & Nolan also observe a tendency for the focused 
constituent to be more likely realized as an \( \iota \)-phrase. In addition, all researchers showed that 
the focused constituent was realized with higher phrase-initial pitch, longer duration, and 
higher intensity. In this sense, Korean overlaps with languages that use stress-based cues to 
mark focus (see section 31.3.1). However, one may interpret these cues as a phonetic effect 
while focus phrasing may be phonological (cf. Jun & Lee 1998).

(5) a. miraneka niel tʃanjake bananaril məknte 
  mira.family.GEN tomorrow night banana.PL eat.PROG 
  ‘Mira’s family is eating bananas tomorrow night.’

b. phrasing in broad focus 

(\( (\text{LH} \ \text{LH})\varphi (\text{LH} \ \text{LH})\varphi (\text{LH} \ \text{LH})\varphi (\text{LH} \ \text{LH})\varphi (\text{LH} \ \text{L})\varphi \text{L%})\iota \)

(\( (\text{LH} \ \text{LH})\varphi (\text{LH} \ \text{LH})\varphi (\text{LH} \ \text{L})\varphi \text{L%})\iota \)

(Korean, Jun & Lee 1998; prosodic phrasing is our own)

For the Bantu language Chichewa, Kanerva (1990) analyses the effect of focus as an insertion 
of a prosodic boundary at the right edge of a focused constituent. Bantu languages are known 
for phrase-penultimate vowel lengthening. In (6a), the penultimate vowel of the phrase-final 
noun \textit{mwála} ‘rock’ undergoes lengthening. In (6), focus within a VP takes the verb as the left 
edge of the prosodic phrase until the right edge of the focused constituent (Kanerva 1990:

(6a) \textit{mwála} nsasiri neimava "he is eating rock"

(6b) \textit{mwála} nsasiri neimva "he is eating rock"
(7c/d). For a more complex analysis of the interaction of focus and prosodic phrasing in Chichewa see Downing et al. (2004) and Downing & Pompolo-Marschall (2013), who however base their analysis on speakers of a different variety of Chichewa.

(6) a. VP focus: What did he do?
(Anaményá nyumbá ndí mwáála)φ
he hit the house with a rock
‘He hit the house with a rock.’

b. PP focus: What did he hit the house with?
(Anaményá nyumbá ndí mwáála)φ

c. Object focus: What did he hit with the rock?
(Anaményá nyuúmba)φ (ndí mwáála)φ

d. Verb focus: What did he do to the house with the rock?
(Anaméenya)φ (nyuúmba)φ (ndí mwáála)φ

(Chichewa, adapted from Kanerva 1990: 156)

Languages thus differ as to where a phrase boundary is inserted: While in Korean a boundary left of the focus is inserted, in Chichewa it is at the right of the focus. Further languages which show similar phrasing patterns under focus are French (Féry 2001; Jun & Fougeron 1995, 2000), Japanese (Beckman & Pierrehumbert 1986; Venditti et al. 2008), Georgian (Skopeteas & Féry submitted), Shingazidja (Patin 2016), and Xhosa (Jokweni 1995; Zerbian 2004).

31.3.3 Register-based cues

Another strategy in a number of languages is that IS affects the pitch register. Pitch register defines reference lines relative to which local tonal targets are scaled (Clements 1979; Ladd 2008). This type of cue is similar to stress-based cues in that it involves increasing the prominence, particularly pitch scaling, of the focused element, and/or it involves the reduction of prominence by compressing the post-focal pitch register; however, this is not achieved through pitch accenting, as these languages generally do not have post-lexical pitch accents. It is also similar to phrase-based cues, in that they also frequently involve pitch scaling, however, in the case of the languages discussed in this section, this is not straightforwardly related to phrasing (though see further section 31.5).

In Mandarin, the pitch register effects under focus are well-studied (Xu 1999; Chen et al. 2016). Two prosodic effects occur, both of which preserve its lexical tones. First, the focused word exhibits a change in f0 that depends on the lexical tone: compared to broad focus, maximum f0 for H tones is raised, including the beginning of HL and the end of LH; while minimum f0 is lowered for L tones, including the end of HL and the beginning of LH. That is, there is a register expansion affecting both the top-line and the bottom-line of the register (Xu 1999:69), in addition to an increasing duration (Chen 2006; Chen & Braun 2006). Second, the f0 after the focused word is lowered, named post-focus compression (PFC) (Xu 1999; 2011; Xu et al. 2012).

Hindi likewise is best characterised as using register-based cues to mark focus prosodically. In this case, there is minimal effect of focal raising, but clear post-focal compression (Patil et al. 2008). In Figure 31.4 compression can be seen on the object in SOV order and the subject in OSV order, with only a small raising on the focused subject in SOV order. The register effect thus appears post-focally and correlates with givenness in Patil et al. (2008). Post-focal compression also functions as a cue to focus perception (Kügler submitted). Other languages that seem best described as register-based are West Greenlandic (Arnhold 2014), Georgian (Skopeteas & Féry submitted), Jaminjung (Simard 2010: 214ff), Serbo-Croatian (with post-
focal compression similar to Hindi) (Godjevac 2005), and Akan, with a general register lowering effect of focus, even for lexical H tones (Kügler & Genzel 2012).

Figure 31.4: Time-normalized pitch tracks in different focus conditions in Hindi, based on five measuring points per constituent, showing the mean across 20 speakers. SOV (left) and OSV word order (right). The comparisons of interest are subject focus (dotted line) and object focus (dashed line) with respect to broad focus (solid line); from Patil et al. (2008: 61).

As for topics, a thorough study on the effects of focus and topic in Mandarin revealed that while topic raises the f0 register at the beginning of the sentence, after the topic f0 drops gradually (Wang & Xu 2011). Hence there is no post-topic compression unlike for foci. The amount of topic raising differs from that of focal raising.

31.4 Syntax-prosody interaction and non-prosodic marking systems

In many languages, syntax is an essential means to encode IS, affecting word ordering and the choice of syntactic construction (see Féry & Ishihara 2016). In many of these languages, there are alignments between prosodic and syntactic encoding of IS, leading to proposals that syntactic encoding of IS is often prosodically motivated. Languages may use a syntactic focus position, most commonly either sentence-initial or -final (e.g. see Rebuschi & Tuller 1999; Drubig & Schaffar 2001; Neeleman et al. 2009; Féry & Ishihara 2016). Sentence-initially, often a specific construction such as a cleft is used (e.g. Lambrecht 2001; Hedberg 2013; Cesare 2014), where the focus in the cleft generally represents the highest prosodic prominence, while the post-focal main clause has reduced prominence (Lambrecht 2001; del Mar Vanrell & Fernández-Soriano 2018).

Word order type of the language and the syntactic focus position appear to covary. In verb-initial languages, an initial focus position seems to be common, and often coincides with the position of nuclear prominence (Herring 1990; Longacre 1995; Simard & Wegener 2017). For example, in Hungarian, narrowly focused items obligatorily appear immediately before the (finite) verb, i.e. initial apart from any preverbal topics (e.g. Kiss 2002; Szendroi 2003). The preverbal position is also the position of the nuclear accent in Hungarian, so the focus must be placed here to align with the nuclear prominence. Narrow focus in Hungarian is marked by increased prosodic prominence on the focused word and/or lowering of accent peaks and deaccenting in the post-focal region, and corrective (exclusive) focus is marked by increasing the relative prominence of the focus compared to the post-focal region (Genzel et al. 2015). Other verb-initial languages reported to have an initial focus position that coincides with the nuclear prominence include Māori (Bauer 1997), Samoan (Calhoun 2015) and probably other Polynesian languages (e.g. Clemens 2014), the Oceanic language Gela (Simard & Wegener 2017), the Berber language Tamasheq (Caron et al. 2015), and the Australian language Dalabon (Fletcher 2014); though see later in this section for counter cases.

Sentence-final focus is common for SVO languages, where the default nuclear prominence position is final, e.g. most Romance languages (Zubizarreta 1998; Ladd 2008; Frota & Prieto 2015a; Büring 2010). For example, in Madrid Spanish foci must be nuclear accented, and the nuclear accent must occur in phrase-final position. The prosodic marking of focus in Madrid Spanish is effectively the same as a standard stress-based system (see section 31.3.1), except that the tendency that the nuclear accent should appear in phrase-final position is weaker than
a tendency against non-canonical word order. Note that there is considerable variation within Romance languages, including between varieties of Spanish, in the extent to which they show in situ stress-based focus marking (Frota & Prieto 2015b; del Mar Vanrell & Fernández-Soriano 2018).

In verb-final languages, the focus position is often immediately before the verb, which correlates with nuclear prominence, with the verb consistently produced with lower pitch, e.g. Hindi (Patil et al. 2008; Féry et al. 2016), Bengali (Hayes & Lahiri 1991), Turkish (Vallduví & Engdahl 1996; Kamali 2011) and Basque (Elordieta & Hualde 2014).

There are a growing number of languages, however, where morphological and/or syntactic focus marking has not been found to correlate with any distinct prosodic marking. For instance, like other Mayan languages the tone language Yucatec Maya has a preverbal focus position, syntactically analysed as a cleft construction (Kügler et al. 2007: 189; Verhoeven & Skopeteas 2015: 3), and canonical word order is VOS. Kügler & Skopeteas (2006) show that the prosody of preverbal focused words does not differ from comparable non-focused words (cf. also Gussenhoven & Teeuw 2008). Moreover, Kügler & Skopeteas (2007) present quantitative evidence that in situ focus on adjectives in an object NP does not affect the tonal realisation either. Similar results as in Yucatec Maya are found for the intonation languages Wolof (Rialland & Robert 2001) and Nleʔkpemxcin (Thompson River Salish) (Koch 2008), and for other African tone languages like Sotho (Zerbian 2007a; Zerbian et al. 2010), Hausa (Hartmann & Zimmermann 2007), Buli (Schwarz 2009), and further ones discussed in Downing & Rialland (2016b), and the Athabaskan languages Beaver (Schwertz 2009) and Navajo (McDonough 2002) as well as in Malay and other varieties of Indonesian (e.g. Maskikit-Essed & Gussenhoven 2016), for which it is shown that neither stress nor prosodic focus is perceptually detected (Goedemans & van Zanten 2007; Roosman 2007).

The initial position is also associated with topics, with many languages placing topic constituents initially which may or may not be integrated syntactically with the rest of the clause (e.g. see Gundel & Fretheim 2004; Neeleman et al. 2009). A commonly found pattern is that initial topics form their own t-phrase, and unlike foci, are not accompanied by any prominence reduction in post-topical material. Languages where this pattern is reported include Hungarian (Surányi et al. 2012; Genzel et al. 2015), German (Féry 2011), Māori (Bauer 1997), Gela (Simard & Wegener 2017), the West African language Zaar, Juba and Tripoli Arabic (Caron et al. 2015), Russian (Alter & Junghanns 2002) and the Australian language Jaminjung (Simard 2010).

### 31.5 Unified accounts

In this chapter, languages are grouped by the main type of prosodic cue each uses to encode IS. However, rather than being independent strategies, it is frequently proposed that these are different instantiations of a general principle of prosodic focus marking (e.g. Truckenbrodt 1995; Zubizarreta 1998; Gussenhoven 2008; Büring 2010; Féry 2013). All three types of cue have been argued to encompass the other two: focus as prominence, focus as alignment and focus as register.

The most commonly proposed unifying principle is prominence, i.e. the smallest prosodic unit (e.g. ω) containing the focus is the most prominent in the largest containing the focus (e.g. t-phrase) (e.g. Truckenbrodt 1995, Samek-Lodovici 2005, Büring 2010). This is consistent with the pre-linguistic effort code (Gussenhoven 2002). Importantly, prominence is an abstract property, so cues may differ between languages. For phrase-based systems, it is argued prominence is at the phrasal level, which can be positional (Büring 2010). For example, as was shown in Korean, φ-phrases to the right of the focus are deleted (section 31.3.2), making
the focused φ-phrase right-most in the τ-phrase, and hence the most prominent (cf. analysis in
Büring 2010). Consistent with this, the focused phrase is phonetically prominent. However,
not all such languages show clear correlates of phonetic prominence of the focus, e.g.
Chichewa (Downing & Pompino-Marschall 2013), weakening this claim. For register-based
systems, pitch range expansion can be argued to mark τ-level prominence. Consistent with
this, in Mandarin Chinese, lexical tones in focused words have more distinct f0 contours, and
less coarticularatory effects (Chen et al. 2016). Syntactic and non-marking languages fit the
principle where the syntactic focus position aligns with the default nuclear stress (see section
31.4). However, when there are no clear cues to nuclear stress, it is hard to see the theory-
external evidence these fit the principle.

Féry (2013) proposes an alternative principle of alignment (see also Koch 2008). She claims
cross-linguistically the strongest tendency is for focused constituents to be aligned with the
left or right edge of a phrase, usually the τ-phrase, or sometimes the φ-phrase; with
prominence being secondary, and separable from alignment. For stress-based systems, she
argues nuclear stress is also phrasal, as it marks a φ-phrase head, therefore, e.g. in Germanic
languages the nuclear accent is right-aligned as it is the right-most phrase head, cf. Figure
31.1, Büring (2010), and Truckenbrodt (1995). However, the phonetic cues to the assumed φ-
phrase boundaries are often weak, e.g. after Peter in Figure 31.1a, weakening this claim. For
syntactic and non-marking languages, similarly to the prominence-led approach, it is difficult
to see the theory-external evidence for alignment where this does not involve phonetic cues to
phrasing. Furthermore, it is not clear in this approach why alignment and prominence so often
coccur, if they are independent.

To our knowledge a fully-fledged focus as register theory has not been advocated for.
However, in current approaches register reference lines are often assumed for languages with
stress-based systems, implying a view of prominence encompassing stress and phrase-based
systems (e.g. German Féry & Kügler 2008; Kügler & Féry 2017; Truckenbrodt 2002). Focal
prominence raises the pitch register line across a phrase: affecting pitch accent height in a
stress-based system, and the whole phrase in a phrase-based system. Féry & Ishihara (2010)
further propose focus raises the register while givenness lowers it; however languages differ
in the extent of raising/compression. For example, post-focal and givenness compression is
almost complete in English and German (Kügler & Féry 2017), but only partial in Mandarin
(Xu 1999) and Hindi (Patil et al. 2008), allowing tonal distinctions to be maintained in
Mandarin.

These unifying principle accounts are appealing and explanatory over a wide range of
languages, however, for all three there remain cases which fit awkwardly at best, especially
for languages without any clear prosodic cues to focus. Rather, these would appear to be
separate, though overlapping approaches. The details of where they overlap, and the extent to
which they do, need much more empirical investigation.

31.6 Evaluation and considerations for future research

From the discussion of the different languages in this chapter, it emerges that certain prosodic
characteristics of a language often entail certain types of prosodic encoding of IS. For
instance, if a language has stress it most likely uses stress-based cues, or if a language uses
predominantly phrase tones to mark intonation units it most likely uses phrase-based cues.
However, this is not without exceptions. Further, as discussed in the previous section, future
work in this area needs to provide more evidence to argue for a prominence view, a phrasing
view or a register view of the expression of focus.
One important topic which we have not had space to cover in this chapter is methodology. Eliciting IS means to elicit both mental states of speakers and hearers and the linguistic means used to convey these mental states. It is not clear that classical tests, i.e. mini-dialogues such as question-answer pairs (cf. Krifka 2008; section 31.2) are sufficient to generate the appropriate mental states; more interactive tasks may be preferred (e.g., *Questionnaire of Information Structure* (QUIS), Skopeteas et al. 2006; Calhoun 2015; Genzel & Kügler 2010; Kügler et al. 2007; Kügler & Genzel 2014; Chen 2018). Elicitation materials need to be carefully constructed, and measured, to control for other effects such as tonal context and segmental influences (cf. Calhoun 2015; Genzel et al. 2015; Wang & Xu 2011; Kügler & Genzel 2012; Féry & Kügler 2008). Further, the majority of studies discussed rely on the acoustic analysis of production data. Only a few studies examine whether listeners perceive and process the prosodic cues according to their pragmatic IS manipulations (e.g. Baumann & Schumacher 2012; Dilley & Heffner 2013; Kügler submitted; Kügler & Gollrad 2015; Ladd & Morton 1997; Rump & Collier 1996; Zerbian 2007a). This is important, e.g. to determine if prominence or phrasing cues are primary.
References


Jun, Sun-Ah & Cécile Fougeron. 1995. The accentual phrase and the prosodic structure of
French. In Kjell Elenius & Peter Branderud (eds.), Proceedings of the 13th International


Dordrecht: Kluwer.

Jun, Sun-Ah & H.S Kim. 2007. VP Focus and narrow focus in Korean. In Jürgen Trouvain &
William J. Barry (eds.), Proceedings of the 16th International Congress of Phonetic
Sciences, 1277–1280. Saarbrücken, Germany, 4-10 August 2007.

Focus in Korean. In Proceedings of fifth International Conference on Spoken Language

Kamali, Beste. 2011. Topics at the PF Interface of Turkish. Massachusetts: Harvard
University Dissertation.

Kanerva, Jonni M. 1990. Focusing on phonological phrases in Chichewa. In Sharon Inkelas &
Draga Zec (eds.), The Phonology-Syntax Connection, 145–161. Chicago: University Of
Chicago Press.

from phonetic prominence in English. Language 87(4). 771–816.


Vancouver, Canada: University of British Columbia.

prominence: Fundamental frequency lends little. The Journal of the Acoustical Society of
America 118(2). 1038–1054.


Kügler, Frank. 2011. The prosodic expression of focus in typologically unrelated languages.
Postdam: Universität Potsdam, Humanwissenschaftliche Fakultät Habilitationsschrift.

Kügler, Frank. Submitted. Post-focal compression as a prosodic cue for focus perception in

Kügler, Frank & Caroline Féry. 2017. Postfocal downstep in German. Language and Speech
60(2). 260–288.

Kügler, Frank & Susanne Genzel. 2012. On the Prosodic Expression of Pragmatic
Prominence: The Case of Pitch Register Lowering in Akan. Language and Speech 55(3).
331–359.

Kügler, Frank & Susanne Genzel. 2014. On the elicitation of focus – prosodic differences as a
function of sentence mode of the context? In Carlos Güssenhoven, Yiya Chen & Dan Dediu
(eds.), TAL-2014: The 4th International Symposium on Tonal Aspects of Languages,

Kügler, Frank & Anja Gollrad. 2015. Production and Perception of Contrast: The case of the
rise-fall contour in German. Frontiers in psychology 6(1254). 1–18.

Kügler, Frank & Stavros Skopeteas. 2006. Interaction of Lexical Tone and Information
Tonal Aspects of Languages, 77–82. La Rochelle.

Kügler, Frank & Stavros Skopeteas. 2007. On the universality of prosodic reflexes of
contrast: The case of Yucatec Maya. In Jürgen Trouvain & William J. Barry (eds.),

Saarbrücken, Germany, 4-10 August 2007.


Skopeteas, Stavros & Caroline Féry. submitted. Focus and intonation in Georgian: constituent structure and prosodic realization.


Figure 31.1: Typical realisations of (1) and (4), showing how focus position affects prosodic realisation. A schematic pitch realization is given, along with the prosodic phrasing, intonational tune, and text, where capitals indicate the pitch accented syllable. See text for further details.
Figure 31.2: Lebanese Arabic (a) and Egyptian Arabic (b) realisation of narrow focus on the initial subject, from Chahal & Hellmuth (2014). As can be seen, post-focal words are deaccented in Lebanese Arabic, but not Egyptian Arabic.
Figure 31.3: Broad focus (a) and contrastive focus (b) in Sardinian, from del Mar Vanrell et al. (2015).
Figure 31.4: Time-normalized pitch tracks in different focus conditions in Hindi, based on five measuring points per constituent, showing the mean across 20 speakers. SOV (left) and OSV word order (right). The comparisons of interest are subject focus (dotted line) and object focus (dashed line) with respect to broad focus (solid line); from Patil et al. (2008: 61).