1aSC36. The "listener" in the modeling of speech prosody. Klaus J. Kohler (Inst. of Phonet. and Digital Speech Processing (IPDS), Univ. of Kiel, D-24098 Kiel, Germany, kjk@ipds.uni-kiel.de)

Autosegmental-metrical modeling of speech prosody is principally speaker-oriented. The production of pitch patterns, in systematic lab speech experiments as well as in spontaneous speech corpora, is analyzed in $f0$ tracings, from which sequences of H(igh) and L(ow) are abstracted. The perceptual relevance of these pitch categories in the transmission from speakers to listeners is largely not conceptualized; thus their modeling in speech communication lacks an essential component. In the metalinguistic task of labeling speech data with the annotation system ToBI, the "listener" plays a subordinate role as well: H and L, being suggestive of signal values, are allocated with reference to $f0$ curves and little or no concern for perceptual classification by the trained labeler. The seriousness of this theoretical gap in the modeling of speech prosody is demonstrated by experimental data concerning $f0$-peak alignment. A number of papers in JASA have dealt with this topic from the point of synchronizing $f0$ with the vocal tract time course in acoustic output. However, perceptual experiments within the Kiel intonation model show that "early," "medial" and "late" peak alignments need to be defined perceptually and that in doing so microprosodic variation has to be filtered out from the surface signal.

1aSC37. Prosodic complexity and phrase length as factors in pause duration. Jelena Krivokapic (Linguist. Dept., Univ. of Southern California, 3601 Watt Way, Grace Ford Salvatori Hall 301, Los Angeles, CA 90089-1693, krivokap@usc.edu)

Research on pauses has mainly focused on predicting the likelihood of pause occurrence and on the effect of syntactic structure on pause duration within an utterance. Very little is known about what factors, apart from syntactic and discourse factors, influence the length of pauses between utterances or phrases. This experiment examines the effect of prosodic structure and phrase length on pause duration. Subjects read 24 English sentences varying along the following parameters: (a) the length in syllables of the intonational phrase preceding and following the pause and (b) the prosodic structure of the intonational phrase preceding and following the pause, specifically whether or not the intonational phrase branches into smaller phrases. In order to minimize variability due to speech rate and individual differences, speakers read sentences synchronously in dyads (Cummins, 2002; Zvonik and Cummins, 2002). The results show that length has a significant effect on pause duration both pre- and postboundary for all dyads, and that prosodic complexity has a significant postboundary effect for some dyads. The possible reasons for the observed pause duration effects and the implications of these results on the question of incrementality in speech production are discussed. [Work supported by NIH DC03172.]

1aSC38. Distinct relative $F0$ levels elicit categorical effects in $F0$ maximum and minimum alignment. Laura C. Dilley and Meredith Brown (MIT Speech Commun. Group, 50 Vassar St., 36-549, Cambridge, MA 02139, dilley@mit.edu)

A standard assumption in intonation research is that the presence and timing of fundamental frequency ($F0$) maxima and minima relative to segments are crucial phonetic characteristics of intonation patterns. The present experiment tests an alternative hypothesis that the representation of intonation patterns is based on the relative pitch levels of syllables in sequence. Synthetic stimuli were created using the phrase *Some lemonade* with an overall rising/falling or falling/rising intonation pattern. Cues to the presence and timing of $F0$ maxima and minima were eliminated by replacing the $F0$ across *lemon-* with level $F0$ contours and replacing the sonorant consonants before and after each target vowel nucleus with Gaussian noise. Four continua were created by shifting the $F0$ levels of one or both syllables in equal 0.5- or 0.75-semitone increments. Thirteen subjects imitated randomized stimuli presented over headphones. Results showed that alignment of maxima and minima in imitations was predictably related to relative $F0$ level: $F0$ maxima (minima) were aligned early in the segmental string when the first target syllable in the stimulus was higher (lower) than the following syllable; otherwise, maxima and minima were aligned late. These results provide support for models of intonation based on relative pitch levels.

1aSC39. Prosodic effects on glide-vowel sequences in three Romance languages. Ioana Chitoran (Dartmouth College Linguist., HB 6087, Hanover, NH 03755)

Glide-vowel sequences occur in many Romance languages. In some they can vary in production, ranging from diphthongal pronunciation [ja,je] to hiatus [ia,ie]. According to native speakers' impressionistic perceptions, Spanish and Romanian both exhibit this variation, but to different degrees. Spanish favors glide-vowel sequences, while Romanian favors hiatus, occasionally resulting in different pronunciations of the same items: Spanish (b[j]ela, ind[j]ana), Romanian (b[i]ela, ind[i]ana). The third language, French, has glide-vowel sequences consistently (b[j]elle). This study tests the effect of position in the word on the acoustic duration of the sequences. Shorter duration indicates diphthong production [jV], while longer duration, hiatus [iV]. Eleven speakers (4 Spanish, 4 Romanian, 3 French), were recorded. Spanish and Romanian showed a word position effect. Word-initial sequences were significantly longer than word-medial ones ($p < 0.001$), consistent with native speakers more frequent description of hiatus word-initially than medially. The effect was not found in French ($p > 0.05$). In the Spanish and Romanian sentences, V in the sequence bears pitch accent, but not in French. It is therefore possible that duration is sensitive not to the presence/absence of the word boundary, but to its position relative to pitch accent. The results suggest that the word position effect is crucially enhanced by pitch accent on V.

1aSC40. The difference between a question and a statement: A cross-dialect survey. Greg Kochanski, Esther Grabe, and John Coleman (Phonet. Lab., Oxford Univ., Oxford, UK)

Seven British English dialects were studied to see what prosodic distinctions are made between statements and questions in read speech. A set of Bayesian classifiers was built upon feature vectors obtained from a spectral analysis of measures of (1) f_0, (2) loudness, (3) spectral slope and (4) voicing periodicity. It was found that the prosodic information useful for the question/statement distinction is distributed broadly across the utterance, and that loudness and spectral slope can be nearly as informative as f_0 (voicing is less informative). The three important acoustic features carry somewhat less than one bit of information each, so prosodic information could be valuable to the listener, and the listener may be able to make the question/statement decision early in the utterance. The contrast differs from one acoustic property to the next: f_0 is marked primarily by slow variations. Conversely, the spectral slope and loudness measurements primarily use shorter-wavelength features, corresponding to structures that are a syllable or two long. We also find substantial differences in the prosodic information that different dialects use, and substantial differences between speakers of the same dialect. [Research supported by the UK Economic and Social Research Council, Grant RES 00-23-1049.]