

Ambiguity resolution has formed a major topic of the course to date. However, a rich source of information that listeners use in resolving ambiguities has been largely omitted both from discussion and in the experiments described. In the majority of the papers on language processing that we have read and discussed up to this point, stimuli have been read rather than spoken. This loss of prosodic information results in sentences that are impoverished relative to those encountered in naturalistic speech processing. The papers assigned for Lecture 12 address this gap by investigating specifically what these prosodic cues are, and how they affect language processing.

Ferreira (2000) identifies three major subclasses of prosodic information – variation in pitch, duration, and stress (~loudness). Each of these may vary independently, and what defines them as prosodic is their additional independence from lexical information. That is not to say that each cannot play a role in distinguishing lexical items – pitch manifests itself in phonemic tone, duration in phonemic segment length, and stress is present on the lexical level, though in well-known cases may shift as a result of phrase-level concatenations (cf. the stress assignment in the phrase “*thirteen Japanese bamboo tables*,” as opposed to that of each word when pronounced in isolation). However, each variable operates on higher-level categories as well. Taken in toto, these cues signal the information structure of utterances and help encode such things as what information is new vs. old or is considered most important.

I am not aware of literature on the reflexes of prosody in languages where all these variables are phonemic. A hypothetical language in which pitch, length and stress all are lexically contrastive seems to offer very little scope for purely prosodic (~phrase-level) information to be encoded. It’s possible that such languages do not exist, however – I believe that languages which have both lexical tone and stress, for example, are rare if not nonexistent. Thus English has lexical stress but not tone (or length), Japanese has lexical tone (pitch-accent; and length) but not stress. In that case there may be a typological generalization that at least one of the variables is set aside for prosody. Note that this is true even though prosodic considerations affect the realization of these lexically-specified variables, as in the English stress example given above and in tone sandhi phenomena in tonal languages (for an example from Chinese, see Shattuck-Hufnagel & Turk 1996).

Shattuck-Hufnagel and Turk (1996) give an overview of a theory of prosodic primitives as well as discussing ways in which these may be represented. In addition, they identify a fourth variable via which prosodic information is encoded, which is the realization of segments. Thus the degree of reduction in vowels, for example, may depend on prosodic factors. While reduction is in some sense an effect of shortened duration, insofar as it also affects vowel quality it has acoustic effects independent of duration. The same holds for consonant lenition (e.g. frication, English flapping) and other prosodic segmental effects.

The prosodic primitives identified are organized hierarchically and range from sublexical categories such as syllables and moras, to the undominated Intonational Phrase. The representational system that is introduced (ToBI) makes use of the markers L and H (low and high, respectively), which are further differentiated by the symbols *, %, and -. This system has the advantage of simplicity and compactness, as well as abstraction away from the acoustic cues of prosodic prominence. This in itself is a major

and necessary attribute of any such system, since as with segmental production extreme acoustic variance and the lack of unvarying cues for any given category is the rule.

Prosody as discussed by these authors is presented as a global phenomenon – structure is assigned to a complete utterance. Of course, this is not how utterances are either produced or perceived, and the papers by Watson and Gibson provide a model for how prosodic representations are built up incrementally. They argue that the structure of these representations depend partly upon the size of the preceding and following constituents, as well as discourse status and a number of other factors relating to syntactic structure, information structure, and more.