

## Prosody and Intonation, Processing Issues

PROSODY AND INTONATION determine much of the form of spoken language. An account of the processing of language—the production and comprehension of words and sentences—must therefore pay attention to prosodic (rhythmic, grouping) and intonational (melodic) structure. The fact that more research in PSYCHOLINGUISTICS has involved written language than spoken language, however, means that the role of prosody and intonation in processing is not yet fully described.

In studies of language comprehension, prosody and intonation have figured in research on SPOKEN WORD RECOGNITION, on SENTENCE PROCESSING (the computation of syntactic structure) and on DISCOURSE processing. One role of prosody and intonation in word recognition is to aid in the operation of segmenting a continuous input into its component words. Studies in many languages (see, for example, the summaries in Otake and Cutler 1996) have shown that listeners can use the rhythmic structure of utterances to determine where word boundaries are most likely to fall. Because rhythmic structure differs across languages, this means that the processes involved in segmenting utterances into words can also be language-specific—stress-based in English (Cutler and Norris 1988), syllable-based in French (Mehler et al. 1981), and mora-based in Japanese (Otake et al. 1993). This language specificity can result in inappropriate application of native-language segmentation procedures to foreign languages with a different rhythmic structure (Otake and Cutler 1996). Young infants can discriminate between rhythmically dissimilar but not rhythmically similar languages (Nazzi, Bertoni, and Mehler 1998).

Whether prosodic and intonational information play a role in the processing of word forms per se—for instance, in the activation of lexical entries—is as yet unresolved. The structure of spoken words again differs across languages in ways that affect this issue. Explicit suprasegmental distinctions between words—for example, TONE in languages such as Thai, pitch accent in languages such as Japanese—constrain word activation and thus show that suprasegmental information can play a role at this level. Suprasegmental cues to LINGUISTIC STRESS in English nevertheless appear to play no part in word activation (Cutler 1986): two words that differ solely in suprasegmental structure, such as *foregoing* (primary stress on the first syllable) and *forgoing* (primary stress on the second syllable), are both activated when listeners hear either one, just as is the case with two words pronounced in every respect identically (such as *sale* and *sail*). However, stress in English is, except in rare pairs such as *foregoing/forgoing*, expressed segmentally (in vowel quality) as well as suprasegmentally, so that segmental information alone may in practice suffice for lexical activation in this language. This is not necessarily the case in other stress languages (Cutler, Dahan, and van Donselaar 1997).

In syntactic processing, the relevant questions have been: do prosody and intonation serve to divide the input into major syntactically motivated chunks? Does such information help to resolve ambiguity, such that sentences which allow more than one interpretation when they are written—for example, *7 read about the repayment with interest*—are effectively unambiguous when spoken? And does prosodic and intonational information determine selection between alternative syntactic analyses that present themselves, albeit temporarily, during the processing even of an unambiguous sentence—for example, between possible continuations of *The horse raced by the—(—gate:—Queen won)!* The evidence is mixed (see special issues of *Language and Cognitive Processes* and *Journal of Psycholinguistic Research* in 1996 for overviews) but in general offers little support for direct availability of syntactic information in prosodic and intonational structure. Prosodic hierarchies, after all, encode specifically prosodic, not syntactic relationships (Shattuck-Hufnagel and Turk 1996; Beckman 1996). Prosody may signal syntactic cohesion (Tyler and Warren 1987), and the presence of a sentence accent or of prosodic correlates of a syntactic boundary can have an effect on syntactic analysis in that it can, for example, lead the listener to prefer analyses that are consistent with the prosody (Nespor and Vogel 1983). But no evidence suggests that syntactic analysis is directly derived from prosodic or intonational cues.

In the comprehension of discourse structure, prosodic salience appears most important; speakers highlight via accent the words that are semantically more central to a message (Bolinger 1978; Ladd 1996), and listeners actively search for accented words because of their central semantic role (Cutler 1982; Sedivy et al. 1995). Furthermore, processing is facilitated by the placement of accent on new information, and the deaccenting of old information (Bock and Mazzella 1983). Experimental evidence suggests that the processing of deaccented words involves integration with an existing discourse model (Fowler and Housum 1987), but it is unclear whether the facilitation observed

with deaccentuation reflects direct exploitation of accent information in discourse-structure decisions or arises indirectly via reference to an existing discourse model in the course of decoding the poorer acoustic information available from deaccented speech. Finally, listeners can interpret prosodic information to derive cues to topic and turn-taking structure in discourse (Hirschberg and Pierrehumbert 1986). Intonational structure is also important for the interpretation of discourse (Pierrehumbert and Hirschberg 1990); the derivation of meaning from intonation involves simultaneous consideration of contours and of the sentential and discourse context in which they appear (Grabe et al. 1997).

The computation of both prosodic and intonational form must of course likewise play a role in speakers' utterance production (Levelt 1989), with prosodic generation referring both to the lexical items and the syntactic structure selected (Ferreira 1993; Meyer 1994), and intonational generation referring to both the sentence to be spoken and its context (Ladd 1996). Production of phonologically alternative forms of words (e.g., via deletion or addition of sounds, as when the middle vowel of *family* is deleted, or a vowel is inserted between the last two sounds of *film*) can be prompted by the prosodic pattern in which the word is uttered (Kuijpers and van Donselaar 1998).

Because, as noted above, there has been less psycholinguistic research on issues specific to spoken language than on the processing of written language, and because it is in addition true that LANGUAGE PRODUCTION has so far attracted far less experimental research than language comprehension, it will be obvious that the production of prosody and intonation is a research area very much in need of wider empirical attention.

See also COMPUTATIONAL PSYCHOLINGUISTICS; INNATENESS OF LANGUAGE; PHONOLOGY; SPEECH PERCEPTION; SYNTAX

—Anne Cutler

#### References

- Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes* 11: 17-67.
- Bock, J. K., and J. R. Mazella. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory and Cognition* 11: 64-76.
- Bolinger, D. L. (1978). Intonation across languages. In J. Greenberg, Ed., *Universals of Human Language*, vol. 2. *Phonology*. Palo Alto, CA: Stanford University Press, pp. 471-524.
- Cutler, A. (1982). Prosody and sentence perception in English. In J. Mehler, E. C. T. Walker, and M. F. Garrett, Eds., *Perspectives on Mental Representation: Experimental and Theoretical Studies of Cognitive Processes and Capacities*. Hillsdale, NJ: Erlbaum, pp. 201-216.
- Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech* 29: 201-220.
- Cutler, A., D. Dahan, and W. van Donselaar. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40: 141-201.
- Cutler, A., and D. G. Norris. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 14: 113-121.
- Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review* 100: 233-253.
- Fowler, C. A., and J. Housum. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26: 489-504.
- Grabe, E., C. Gussenhoven, J. Haan, E. Marsi, and B. Post. (1998). Preaccental pitch and speaker attitude in Dutch. *Language and Speech* 41, 63-85.
- Hirschberg, J., and J. Pierrehumbert. (1986). The intonational structuring of discourse. *Proceedings of Twentyfourth Association Computational Linguistics* 134-144.
- Kuijpers, C., and W. van Donselaar. (1998). The influence of rhythmic context on schwa epenthesis and schwa deletion. *Language and Speech* 41: 87-108.
- Ladd, D. R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Mehler, J., J.-Y. Dommergues, U. Frauenfelder, and J. Segui. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior* 20: 298-305.
- Meyer, A. S. (1994). Timing in sentence production. *Journal of Memory and Language* 33: 471-492.
- Nazzi, T., J. Bertoni, and J. Mehler. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24: 756-766.
- Nespor, M., and I. Vogel. (1983). Prosodic structure above the word. In A. Cutler and D. R. Ladd, Eds., *Prosody: Models and Measurements*. Heidelberg: Springer, pp. 123-140.
- Otake, T. and A. Cutler, Eds. (1996). *Phonological Structure and Language Processing: Cross-Linguistic Studies*. Berlin: Mouton.
- Otake, T. G. Hatano, A. Cutler, and J. Mehler. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language* 32: 358-378.
- Pierrehumbert, J., and J. Hirschberg. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, and M. E. Pollack, Eds., *Intentions in Communication*. Cambridge, MA: MIT Press, pp. 271-323.
- Sedivy, J., M. Tanenhaus, M. Spivey-Knowlton, K. Eberhard, and G. Carlson. (1995). Using intonationally marked presuppositional information in on-line language processing: Evidence from eye movements to a visual model. *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, pp. 375-380.
- Shatuck-Hufnagel, S., and A. E. Turk. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25: 193-247.
- Tyler, L. K., and P. Warren. (1987). Local and global structure in spoken language comprehension. *Journal of Memory and Language* 26: 638-657.

#### Further Readings

- Friederici, A., Ed. (1998). *Language Comprehension: A Biological Perspective*. Heidelberg: Springer.
- Journal of Psycholinguistic Research*. (1996). *Special Issue on Prosodic Effects on Parsing* 25(2).
- Language and Cognitive Processes*. (1996). *Special Issue on Prosody and Parsing* 11(Vi).