Pitch accent type matters for online processing of information status: evidence from synthetic speech and human speech

Aoju Chen (aoju.chen@mpi.nl)
Max Planck Institute for Psycholinguistics

It is well established in (psycho)linguistic literature (Cutler, Dahan, van Doneselaar 1997, Gussenhoven forthcoming) that in Germanic languages, placement of pitch accent is crucial for interpreting information status. In contrast, the role of pitch accent type is far from clear. Theories of intonational meaning (Brazil 1975, Gussenhoven 1984, Pierrehumbert & Hirschberg 1990, Steedman 2000) claim that different types of pitch accent convey different types of information status. However, there is no consensus on the postulated meanings of the pitch accents. Empirically, the relation between pitch accent type and information status is largely unexplored. Exceptional are Baumann and Hadelich (2003), who studied H* and H+L* in marking givenness in German, and Watson, Tanenhaus, Gunlogson (submitted), who investigated H* and L+H* in marking newness and contrastive givenness in American English.

Adopting the eye-tracking paradigm used in Dahan, Tanenhaus and Chambers (2002), the present study set out to pin down the role of H*L, L*H L, H*L H and L*H, in interpreting give. vs. new information in both human and synthetic speech in British English. Eye fixations were monitored as participants followed pre-recorded instructions to move objects displayed on a computer screen via a computer mouse. On each experimental trial, two of the four objects displayed had names that were phonemically similar (e.g., *comb, coat*). Each trial consisted of two consecutive instructions (e.g., *Put the comb/coat below the triangle; now put the comb below the triangle).* The object mentioned in the 2nd instruction was the target; its phonemically related counterpart served as the competitor. The 1st instruction mentioned either the target or the competitor, marking the competitor either new or given during the 2nd instruction. The target word in the 2nd instruction was said with each of the four pitch accents as well as deaccentuation. The instructions in human voice were recorded by a male speaker in Southern Standard British English. The instructions in synthetic voice were generated with the Festival diphone synthesiser.

The data of each participant's right eye were coded in terms of fixations from the onset of the target word in the 2nd instruction to the moment when participants clicked on the target picture. Proportion of fixations (PoF) to the competitor picture was calculated in 33-ms time interval for each condition and each participant. It was found in human speech that when the competitor was given, L*H and deaccentuation triggered a higher PoF to the competitor picture than H*L and L*H L; when the competitor was new, H*L and L*H L triggered a higher PoF to the competitor picture than L*H and deaccentuation. These results clearly show that pitch accent type matters when interpreting information status, with the falling accents creating a bias towards newness and the rise accent, like deaccentuation, creating a bias towards givenness. Interestingly, the effect of pitch accent type is also reflected in how fast the participants could adjust their decision as to which picture to move before the name of the picture was fully revealed. For example, when the competitor was given, the PoF to the competitor picture increased in most accent conditions in the first 300 ms as a result of a bias towards a given entity, but started to decrease substantially earlier in the H*L condition than in the L*H and deaccentuation. Remarkably, the same effects were observed when listeners were presented with synthetic speech of which the segmental quality may not be ideal. As regards H*L H, listeners appeared to be biased to the target picture independent of what was mentioned before. An explanation for this unexpected finding is proposed in the light of the effect that the duration of phonemically identical sequences can have on lexical recognition. In addition, topics for further investigation are suggested.

References

Baumann, S., and Hadelich, K. (2003). Accent type and givenness: an experiment with auditory and visual priming. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1811-1814). Barcelona: Universitat Autònoma de Barcelona.

Brazil, D. (1975). *Discourse intonation*, vol. 1. Birmingham University, Department of English.

Cutler, A., Dahan, D., and van Doneselaar. W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech*, 40, 141-201.

Dahan, D., Tanenhaus, M.K., and Chambers, C.G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47, 292-314.

Gussenhoven, C. (1984). *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

Gussenhoven, C. (forthcoming). Types of Focus in English. In D. BÄuring, M. Gordon, and C. Lee (eds.) *Topic and Focus: Intonation and Meaning. Theoretical and Crosslinguistic Perspectives.* Dordrecht: Kluwer.

Pierrehumbert, J. B. and Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In R. R. Cohen, J. Morgen, and M. E. Pollack (Eds.), *Intentions in Communication* (pp. 271-311). MA: MIT press.

Steedman, M. (2000). Information struction and the syntax-phonology interface. *Linguistic inquiry*, 31 (4), 649-689.

Watson, D., Tanenhaus, M.K., and Gunlogson, C.A. (submitted). Interpreting pitch accents in on-line comprehension. Retrieved from http://www.bcs.rochester.edu/people/dwatson/#publications