The Realisation of Intonational Plateaux: Effects of Foot Structure[*]

Rachael-Anne Knight

Department of Linguistics, University of Cambridge

There is a general acknowledgement in the intonation literature that tones are often not realised as single turning points in the fundamental frequency contour, frequently occurring more as flat plateaux. House *et al.'s* (1999) study of a single-speaker database demonstrates that the realisation of these plateaux co-varies with linguistic structure. The present study extends House *et al.'s* analysis of the realisation of the plateau to five additional speakers. The earlier results are largely replicated and extended to cover an additional prosodic structure and to take account of additional measurements. Results indicate that differences found for proportional duration of the plateau are largely due to the effects of foot structure on syllable duration rather than on the absolute duration of the plateau. In addition it is suggested that plateau alignment may mark linguistic structure, specifically the difference between mono- and polysyllabic feet.

1   INTRODUCTION

Many theories of intonational phonology, especially those in the autosegmental-metrical tradition popularised by Pierrehumbert (1980), represent the intonation contour as a string of high (H) and low (L) targets. These targets are considered to be the linguistically important points of the contour whilst the transitions between them are seen as linguistically unimportant interpolations. In recent years much work has focused on describing how targets are aligned with the segmental string. In such investigations it is common to visually identify turning points associated with high and low targets and express their alignment as a duration or a percentage into some tone-bearing unit such as the syllable or foot.

Despite this general method it is often acknowledged in the literature that the search for turning points is by no means an easy one. There are many factors that make the process difficult. For example, turning points may be obscured by voiceless sections or microprosodic perturbations in the contour. Even when these are minimal, such as in sonorant stretches, the turning points may still be hard to locate as high and low parts of the contour may be sustained forming plateaux rather than sharp peaks and troughs. This phenomenon is very common (e.g. Silverman and Pierrehumbert, 1990, House and Wichmann, 1996, d'Imperio, 2002) and raises the question of whether a particular point in the plateau (such as the beginning, middle or end) or the whole plateau itself should be considered to be the speaker's real target.

When writing rules for speech synthesis this question is crucial if contours are to be expressed as a fixed number of turning points aligned with the segmental string. House *et al.* (1999) studied a medium-sized database from one male speaker of Southern British English. They attempted to reduce natural falling nuclear (H*L) contours to a small number of turning points sufficient for synthesis. Visual analysis revealed that the high tone was often realised as a plateau and informal auditory testing indicated that both ends of these plateaux needed to be represented for natural sounding synthesis to result. Therefore, they decided to use