

On the Perception of Juncture in English

W.J. Barry

Kiel, Federal Republic of Germany

1. Introduction

In English, where relatively unconstrained syllabification and free stress patterning exist, word-boundary disambiguation is more critical than in languages where fixed word stress or more restricted syllable structure reduce potential ambiguities to a minimum. Word identity is, of course, guaranteed to a large extent by situational information and by on-going processing of syntactic and semantic structure (Marslen-Wilson, 1975; 1980). However, it may be presumed in this as in other areas of speech perception that top-down processing must be supported by a systematic signal structure.

In the two major experimental studies of juncture production and perception (Lehiste, 1960; Gårding, 1965), systematic differences in the speech signal structure of juncturally ambiguous strings of segments were found. With few exceptions (e.g. the style-dependent use of glottalized vocalic onset in initial vowels) the differences can be subsumed under the parameters duration and intensity, both of which contribute to the perceptual prominence of a segment. Word initial segments generally have greater prominence than cognate word-final segments. In a later study by the present author (Barry, 1981), the main trends of the previous analyses were confirmed, though there was a tendency, to a differing degree from speaker to speaker, to neutralize word-initial and word-final parameter values in running texts. Also, uncertainty in junctural identification was found to coincide with one or more atypical parameter values.

Neither the studies mentioned above nor the exclusively perception oriented studies by O'Connor and Tooley (1964) and by O'Connor-Dukes and Nakatani (1979) have carried out perception tests with systematic variation of the signal properties which have been found to differ.

The present paper reports on a perception experiment which aimed to examine the contribution to junctural identity of various parameters specified in the three studies mentioned above.

The junctural pairs: See Mabel - seem able; why choose - white shoes; keep sticking - keeps ticking were spoken several times by the author and one pair selected on the basis of an auditory judgement as to junctural distinctiveness and prosodic identity. Care was taken not to produce glottal constriction during 'seem able', nor to glottalize the /t/ in 'white shoes'. These pairs were

selected from the nine pairs used in the previous study (Barry, 1981) for reasons of differing trans-junctural complexity. The durational characteristics of the individual utterances are given in Table I.

Apart from the modified properties, no changes were made in the stimuli. Table II gives the parameter values used for modifying the stimuli.

Signal manipulation was carried out with the Speech Signal Processor (Schäfer, 1982) at the Kiel Institute of Phonetics. To avoid undue concentration on the acoustic structure of each juncture type, all three pairs were offered together in a single test.

Five repetitions of the 32 generated stimuli from the 3 junctural pairs (6 + 8 + 18) were offered with a 3 s. decision pause in randomized order to 20 native speakers of English for identification. Subjects were required to mark the member of the pair identified.

Table I. Durational values for signal segments of junctural pairs (ms)

a.	/s/	/i:/	/m/	ei/	/b/	/l/	Σ
See Mable	153	149	76	187	75	157	797 ms
Seem able	144	161	61	187	82	135	770 ms
b.	/waɪ/	/t/	/ʃ/	/u:z/	Σ		
Why choose	223	57	110	438	828 ms		
White shoes	218	54	174	404	850 ms		
c.	/ki:/	/p/	/s/	[t]	[-h]	/ɪkɪp/	Σ
Keep sticking	139	102	95	28	16	335	715 ms
Keeps ticking	124	80	81	56	39	334	714 ms

Table II. Parameter values for modification of junctural-pair members

a.					
See Mabel vs. seem able	59	67,6	76		[m] duration (ms)
step	1	2	3		
b.					
why choose vs. white shoes	174	152	131	110	[ʃ] duration (ms)
step	1	2	3	4	
c.					
keep sticking vs. keeps ticking	161	179	197		[ps] duration (ms)
step	1	2	3		[t ^h] duration (ms)

2. Results

Group average scores are displayed graphically in Fig. 1a - c. It is immediately apparent that the degree of certainty with which the junctural distinction is made varies considerably. Although the number of 'Mabel'-judgements increases systematically with increasing [m]-duration (all rank-sum diffs. > 12, n = 20, k = 3; p < 0.05 after Wilcoxon-Wilcox, one-sided), even the unmodified, original 'Mabel' does not score better than 60%. The 'choose' judgements, on the other hand, range from 24% - 93%, and 'sticking' from 10% - 88%.

These differences can be explained partly by the degree of acoustic change imposed on the original stimuli (cp. Table II). Whereas the difference in [m]-duration between 'See Mabel' and 'seem able' is only 17 ms, the fricative element in 'white shoes' and 'why choose' differs by 64 ms. However, the even greater difference between 'keep sticking' and 'keeps ticking' (36 ms + 51 ms) does not result in the expected unanimity of judgement for the original stimuli.

With the exception of the 'sticking-ticking' pair the parameters modified in this experiment appear to explain the junctural distinction completely. In neither of the other two pairs is there any difference in the judgements as a function of the original stimulus. The stimuli derived from original 'ticking', however, receive significantly more 'ticking' judgements than those derived from 'sticking' (T = 32,5 < 52, n = 20, p < 0.05 after Wilcoxon). This indicates that remnant junctural information is contained in the stimuli apart

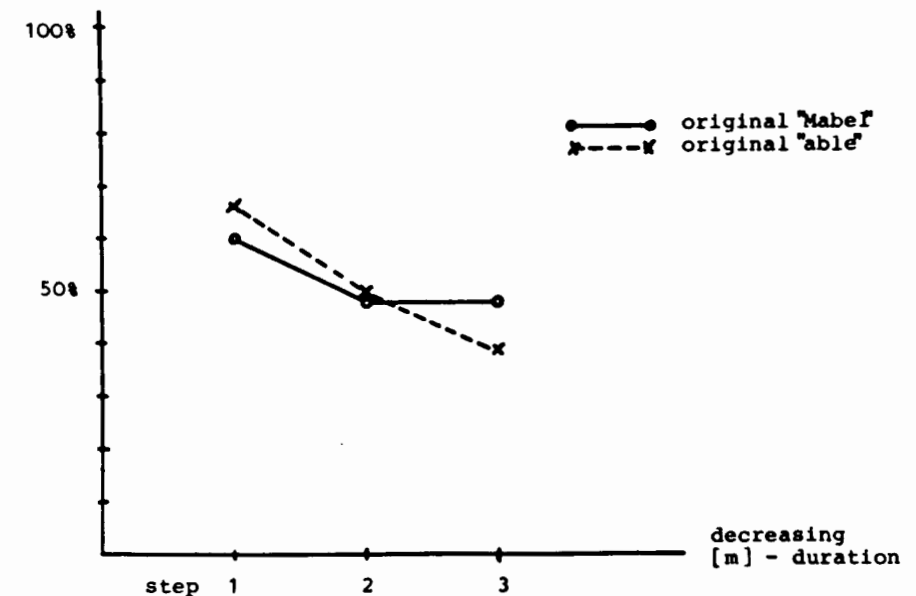


Figure 1. 'Mabel' judgments (English subjects).

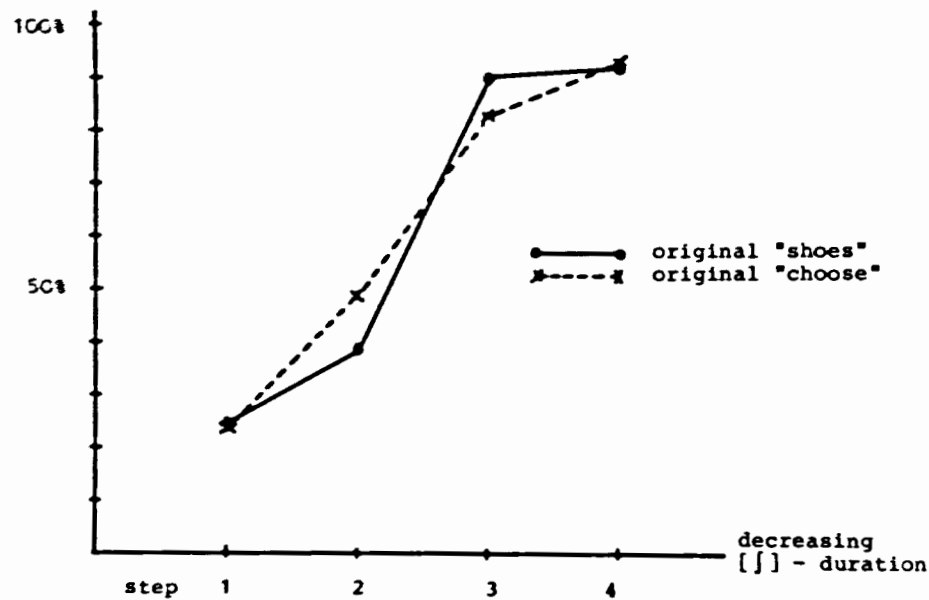


Figure 1b. 'choose' judgments (English subjects).

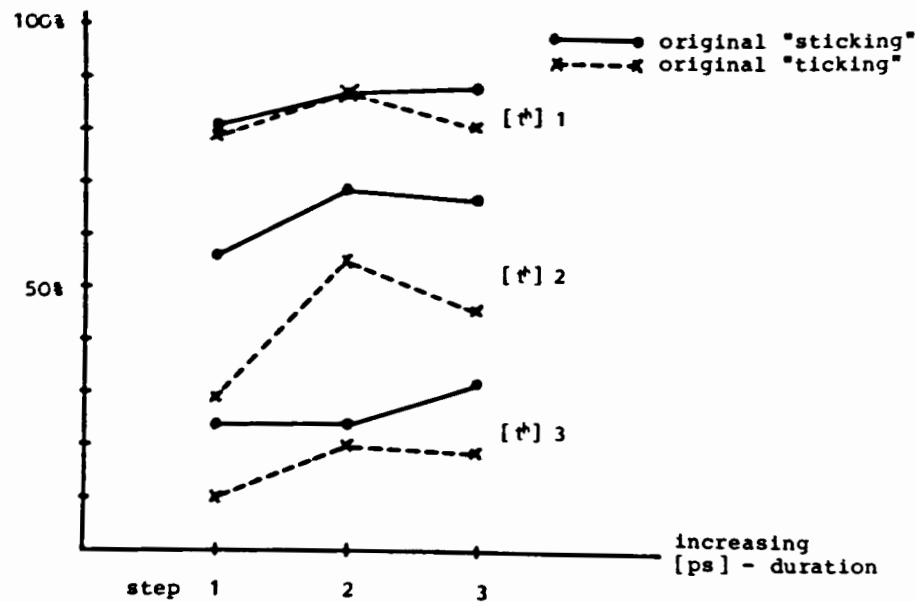


Figure 1c. 'Sticking' judgments (English subjects).

from the durational information in the [pst^h] segments that were manipulated. The likeliest candidate is the greater release burst intensity of [t^h] since it remained unmodified by the durational manipulation applied to the section following the burst spike. This assumption is plausible in the light of the dominant influence which the aspiration has on junctural identity: changes in /ps/ duration have only a slight effect on judgement distribution (see Fig. 1c) though the influence again appears to be systematic between the extreme parameter values (diff. > 12, n = 20, k = 3; p < 0.05 after Wilcoxon-Wilcox).

In a test with 15 German subjects the regularities in perception were much less clearly defined (Fig. 2a - c), though for the 'choose-shoes' and the 'ticking-sticking' distinction there was a significant perceptual effect of [ʃ] and [t^h]-duration, respectively (difs. > 15 for 'shoes-choose', k = 4; diffs. > 11 for 'sticking-ticking', k = 3; p < 0.05). The German subjects were unable to distinguish the 'Mabel-able' pair, understandably in view of the predominance of glottalized vowel onsets in German, and they failed to react to the [ps]-duration in 'keeps ticking' vs. 'keep sticking'.

3. Discussion

The following tentative conclusions may be drawn from the results:

1. Junctural distinctions vary considerably in their perceptibility, but under circumstances of mixed presentation do not achieve scores suggesting clear junctural categories even in their original realisation. This would suggest that the *phonetic* decoding of word boundaries is secondary to the

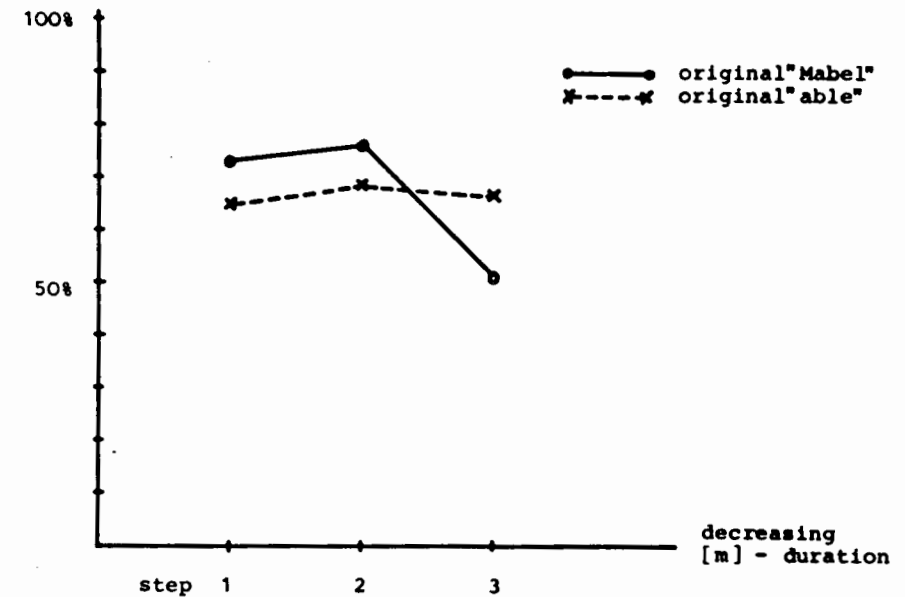


Figure 2a. 'Mabel' judgments (German subjects).

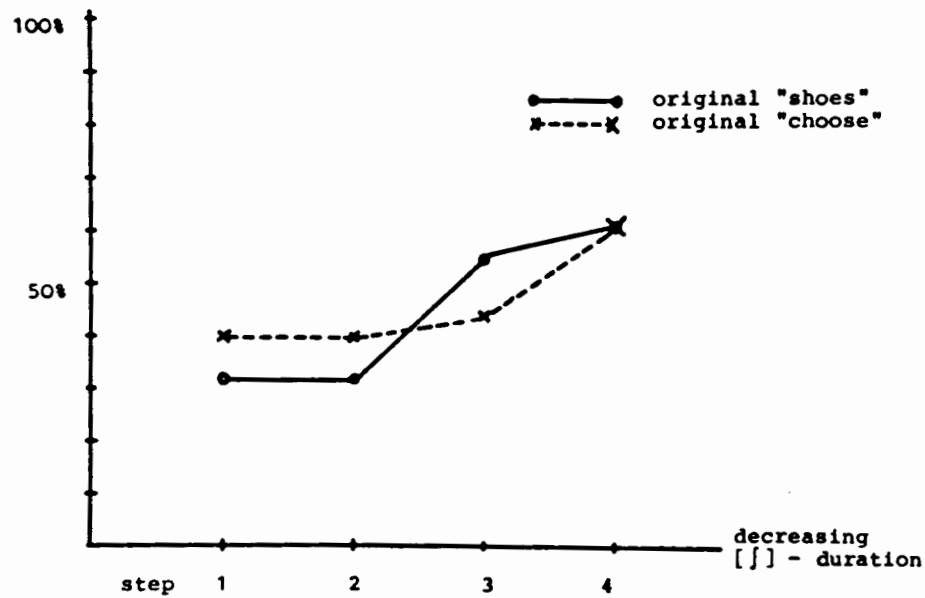


Figure 2b. 'choose' judgments (German subjects).

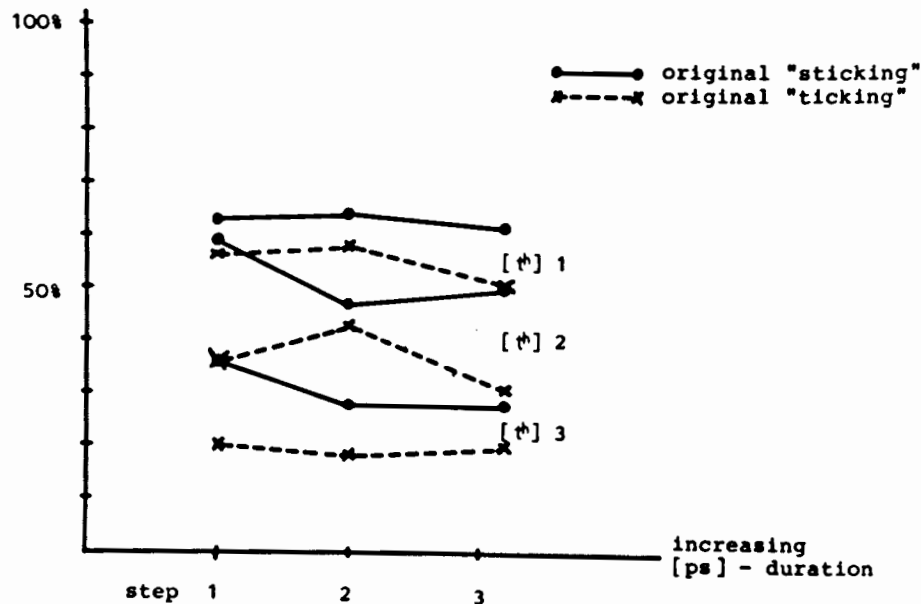


Figure 2c. 'sticking' judgments (German subjects).

clearly categorical function of phonemic identification, and that top-down information has a more prominent place in the perceptual strategy. This is in accord with the tendency to neutralize junctural distinctions in running text *production* (Gårding, 1965; Barry, 1981).

- Not all systematic signal differences contribute to the perception of juncture. There is some support for Gårding's (1965, p. 15) hypothesis that *post*-junctural information is more important than *pre*-junctural information. The degree of post-junctural aspiration was dominant in the 'ticking-sticking' decisions, and although the 'choose-shoes' pair was only varied in a single (post-junctural) parameter it was distinguished quite as well as 'sticking-ticking'. Also it may be assumed that the missing [ʔ] in 'able' contributed to the extreme uncertainty of decisions on the 'Mable-able' pair. However, a global post- vs. pre-junctural dichotomy is obviously too undifferentiated; the longer post-junctural /s/ in 'sticking' had no effect on decisions.
- With mixed presentation (i.e., under slightly less 'clinical' conditions than often exist in perception tests), junctural identification scores could be explained completely by a single parameter. This suggests that subjects' extreme sensitivity to a multiplicity of signal differences, which has been used as an argument against feature-based perception (Bailey and Summerfield, 1978), might be less a component of normal speech perception than a sign of human perceptual flexibility and learning potential under extreme condition (cp. Barry, 1980, p. 116-7).
- Despite the 'production-universal' character of the junctural differences, the perceptual exploitation of them appears, at least partly, to be language specific. The German subjects, accustomed to a language with strong distributional constraints in the phonemic system (final devoicing, glottal onset in initial vowels) showed generally weaker reactions to the potential junctural signals.

References

- Bailey, P.J. and Summerfield, A.Q. (1978). Some observations on the perception of [s]+stop clusters. *Haskins Labs. Status Report on Speech Research*, SR-53.2, 25-60.
- Barry, W.J. (1980). *Die Verarbeitung akustischer Information in der lautsprachlichen Wahrnehmung*. (= Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK), 13).
- Barry, W.J. (1981). Internal juncture and speech communication. In: *Beiträge zur experimentellen und angewandten Phonetik* (eds. W.J. Barry and K.J. Kohler) (= Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK), 16), 229-289.
- Gårding E. (1965). *Internal Juncture in Swedish*. Lund: Gleerup (= Travaux de l'Institut de Phonétique de Lund VI).
- Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. *Phonetica* 5, Supplement 5-54.
- Marslen-Wilson, W. (1975). Sentence perception as an interactive parallel process. *Science* 189, 226-228.
- Marslen-Wilson, W. (1980). Speech understanding as a psychological process. In: *Spoken Language Generation and Understanding* (ed. J.C. Simon), Dordrecht, Boston, London: Reidel Publ. Co., 39-67.

- O'Connor, J.D. and Tooley, O.M. (1964). The perceptibility of certain word boundaries. In: *In Honour of Daniel Jones* (eds. Abercrombie, Fry, MacCarthy, Scott and Trim), London: Longmans, 171-176.
- O'Connor-Dukes, K.D. and Nakatani, L.H. (1979). The role of morphology in a word-boundary decision task. *J. Acoust. Soc. Am.* **66**, Suppl. 1, 550 (A).
- Schäfer, K. (1982). Concepts in the SSP programme. In: *Phonetic Data Processing at Kiel University* (eds. W.J. Barry and K.J. Kohler), (=Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK), **18**, 110-121.