View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Leiden University Scholary Publications

Effects of focus distribution and accentuation on the temporal and melodic organisation of word groups in Dutch

Vincent J. van Heuven
HIL/Leiden University

1. Introduction[1]

A speaker speaks in order to be heard in order to be understood. The words in an utterance differ in their communicative importance. Some words merely recapitulate predictable referents that were introduced earlier on in the discourse ('old' or [–focus] elements); others introduce highly informative, contextually and/or situationally rather unpredictable, referents ('new' or [+focus] elements). A speaker, governed by the principle of least effort, will speak as quickly and as sloppily as the hearer will let him get away with. In order to reduce articulatory effort and at the same time maintain the required minimum of perceptual distinctivity, the speaker strikes a compromise: he will pronounce 'new' [+focus] elements relatively carefully, slowly and deliberately (*hyper-articulated*); 'old' [–focus] materials, however, will typically be slurred (*hypo-articulated*).

There is a sizable literature to support this specific prediction of Lindblom's (1990) 'hyper and hypo' theory (or H&H theory). Lieberman (1963) showed, for instance, that the word *nine* in the well-known English expression *A stitch in time saves nine* is pronounced more sloppily and with less intensity than the same word in the less pregnant context *The next number is nine*. Moreover, a perceptual experiment indicated that the hyper-articulated word, when excised from its original spoken context, was more intelligible than the hypo-articulated counterpart. These effects were replicated on a larger scale by Hunnicut (1985). Similarly, Fowler & Housum (1987) showed that repeated referents introduced in the preceding sentence (and which were therefore [–focus]), were spoken more quickly than the same words when these were newly introduced into the discourse (and therefore [+focus]).

The prosodic head of a [+focus] word or word group (henceforth 'constituent') in languages such as Dutch and English is marked by a pitch accent, i.e., has a conspicious change in pitch associated with the stressed syllable. A secondary effect of accentuation is that the word with the pitch change associated to it, is expanded in time by some 10%; moreover, the time-expansion is not limited to just the stressed syllable, but affects unstressed syllables as well (Nooteboom, 1972; Eefting, 1991; Sluijter & van Heuven, 1995, 1996). It is unclear, therefore, whether the temporal expansion of 'old' referents in Fowler & Housum's study is the result of their being [+focus] or of their being pitch accented.

The first to tackle this question was Eefting (1991), whose experimental materials were constructed so as to disentangle the effects on temporal expansion due to focus structure on the one hand and of accentedness on the other. Her results show that only the word that carries the pitch accent is temporally expanded; if a [+focus] word remains unaccented (which may

[1] This squib is based on a lecture presented at the Linguistics in the Netherlands meeting, on January 21, 1995 in Utrecht. The experiment was run by Esther Hofman and Clyde Joseph as a course requirement.

happen when the word is not the prosodic head of its larger constituent), it is not lengthened. Unfortunately, Eefting's experiment may not be conclusive, on two counts.

The first problem with Eefing's materials is that the [+accent, +focus], [–accent, +focus] and [–accent, –focus] words were compared across different sentences. Though the target words themselves were identical, their temporal organisation may therefore differ due to uncontrolled contextual effects. Consider, as a case in point, the target word *Kees* (proper name). This target occurred in the complex, coordinated noun phrase *Kees ten Kate en Marie van der Bilt* in the comparison between [+focus, +accent] (6th of 25 syllables in sentence) versus [+focus, –accent] (5th of 19 syllables); it occurred in the complex non-coordinated noun phrase *Kees ten Kate* in the comparison of [–focus, +accent] and [–focus, –accent] (5th of 18 syllables in both conditions), and finally it occurred just by itself in sentences pair comparing [+focus, +accent] and [–focus, –accent] (5th of 15 syllables in both conditions). Dutch words are typically time-compressed as they occur earlier and/or in longer sentences (de Rooij, 1979; Rietveld & van Heuven, 1997: 285-286).

The present experiment was set up to check the basic validity of Eefting's conclusions, using properly controlled materials. Eefting's recordings were made of a single male speaker, who may not be representative of the Dutch linguistic community at large. It was shown, for instance, that – unlike the Dutch population at large – this particular talker, a well-known newscaster and speaker of commercials, exhibits no spectral vowel reduction in unstressed full vowels (van Son & Pols, 1990). Therefore, our experiment, in contrast, used a smaller set of sentence types but collected these from a larger set of speakers, equally divided over the sexes. Finally, we examined the effects of a wider range of focus conditions (see below), not only on the temporal organisation of the target constituents but also on their melodic realisation.

2. Method

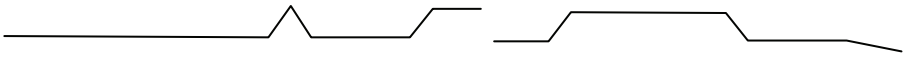
Two lexically different sentences were constructed, with the target constituent in the object position. In the short sentence the target was a simple Det+Adj+Noun NP; here the prosodic head (PH)is the noun, whilst the adjective is the prosodic dependent (PD). The target constituent in the longer sentence was an NP (PD) followed by a PP (PH).

		PD	PH	
Karel	*heeft*	*[een paarse*	*papegaai]*	*gezien*
'Charles	has	[a purple	parrot]	seen'
Olivier	*heeft*	*[het apartement*	*op de heuvel]*	*gekocht*
'Oliver	has	[the apartment	on the hill]	bought'

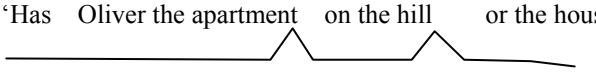
These two sentences were preceded by six different questions, manipulating the focus structure of the following answer such that the target constituent would be produced with double focus (on both PD and PH), narrow focus on either PD or PH, no focus on either, with the target constituent once in prenuclear and once in postnuclear position. The crucial condition contained PD and PH in a single integrative focus domain. Here the [+focus] status of the entire domain is signalled by a single 'integrative' accent on the PH, so that PD is [+focus] and yet [–accent]. If [+focus] information is indeed communicatively important, the H&H hypothesis predicts that it should be lengthened by the speaker. If, on the other hand, the

lengthening effect reported in the literature is merely a concomitant property of a pitch accent on a word, no lengthening of the PD should obtain. The six focus conditions are illustrated below. Accented syllables are printed in small caps; the [+focus] material in the answer sentences appears in square brackets. Stylized intonation patterns have been drawn over both question and answers.

- **Double focus**, i.e. accent on PD and PH: (Double)

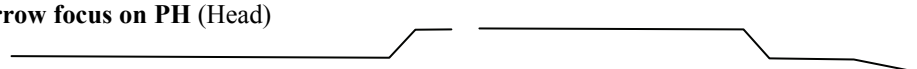


Q: *Heeft Olivier het aparteMENT op de HEUvel of het HUIS in het DAL gekocht?*
'Has Oliver the apartment on the hill or the house in the valley bought?'

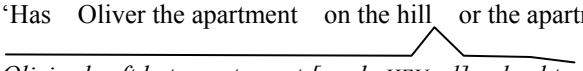


A: *Olivier heeft [het aparteMENT] [op de HEUvel] gekocht.*

- **Narrow focus on PH** (Head)

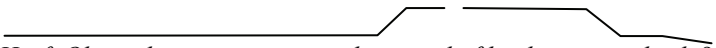


Q: *Heeft Olivier het apartement op de HEUvel of het apartement in het DAL gekocht?*
'Has Oliver the apartment on the hill or the apartment in the valley bought?'

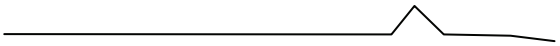


A: *Olivier heeft het apartement [op de HEUvel] gekocht.*

- **Integrative focus**, i.e. no accent on PD but accent on PH (Integrative)

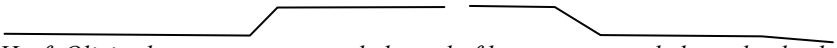


Q: *Heeft Olivier het apartement op de HEUvel of het kasTEEL gekocht?*
'Has Oliver the apartment on the hill or the castle bought?'

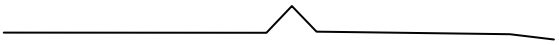


A: *Olivier heeft [het apartement op de HEUvel] gekocht.*

- **Narrow focus on PD** (Dependent)

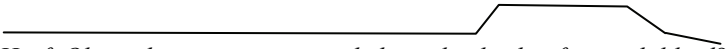


Q: *Heeft Olivier het aparteMENT op de heuvel of het HUIS op de heuvel gekocht?*
'Has Oliver the apartment on the hill or the house on the hill bought?'

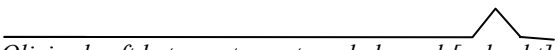


A: *Olivier heeft [het aparteMENT] op de heuvel gekocht.*

- **No focus, target in prenuclear position (Prenuclear)**



Q: *Heeft Olivier het apartement op de heuvel gekocht of geschilderd?*
'Has Oliver the apartment on the hill bought or painted?'



A: *Olivier heeft het apartement op de heuvel [gekocht].*

- **No focus, target in postnuclear (Postnuclear)**

  Q: *Heeft Olivier of Adriaan het apartement op de heuvel gekocht?*
  'Has Oliver or Adrian the apartment on the hill bought?'

  A: *[Olivier] heeft het apartement op de heuvel gekocht.*

Two male and two female native speakers of standard Dutch, language students at Leiden University but naive as to the purpose of the experiment, read the set of question-and-answer sequences twice. These speakers were selected post hoc from a larger group of subjects on the grounds that only they had faithfully copied the speech melody of the model versions of the answer sentences (see next paragraph).

The speakers were seated in a sound-insulated recording studio, and had the texts in print before them. They listened over headphones to the questions and answerd which were generated in real time using diphone synthesis (van Rijnsoever, 1989) with standardized intonation patterns (as indicated in the examples above) but with all duration rules blocked. The pitch accents in the answer sentences were realised as rise-fall configurations ('1&A' or 'pointed hat' in 't Hart, Collier & Cohen, 1990; H*L in autosegmental notation, cf. Rietveld & van Heuven, 1997: 263-277). As a result, the synthesized patterns reflected the presence or absence of pitch accents, in accordance with the focus structure, but not the temporal effects. Subjects listened to each question-answer pair in turn and immediately repeated the answer sentence, mimicking its sentence melody to the best of their ability. Recordings were made through a Sennheiser MKH416 condenser microphone onto digital audio tape, and stored on computer disk (downsampled to 16 kHz, 16 bits). Fundamental frequency (F_0) curves were determined by the method of subharmonic summation (Hermes, 1988). The curves were corrected when necessary, and syllable durations determined, by hand using the GIPOS speech processing software.

3. Results and conclusions

For both the prosodic head and for the dependent the F_0 rise and fall were stylized by hand. The excursion size in semitones Hz (st) of the rise-fall configuration was equated with that of either the rise or the fall portion, whichever was the larger of the two. Figure 1 plots mean excursion size of head and dependent for each of the six focus distributions.

Figure 1 shows a relative large pitch movement (>5 st) on the prosodic head when the focus distribution predicts an accent on it, i.e,. in double focus, in narrow focus on the head, and with integrative focus. The pitch excursion on the head is relatively small (<5 st) when it is outside focus. Conversely, when the dependent is in narrow focus, or part of a double focus it bears a large pitch movement (>5 st); the excursion size is relative small (<4 st) when the focus distribution predicts the absence of an accent on the dependent. Generally, the intended focus distribution can be reconstructed perfectly from the excursion sizes of the pitch movements on head and dependent. The notable exception is the non-distinctness of narrow focus on the head versus integrative focus; this, however, is precisely in line with the literature,

where a pitch accent on just the prosodic head is claimed to completely ambiguous between narrow and integrative focus (van Heuven, 1994).

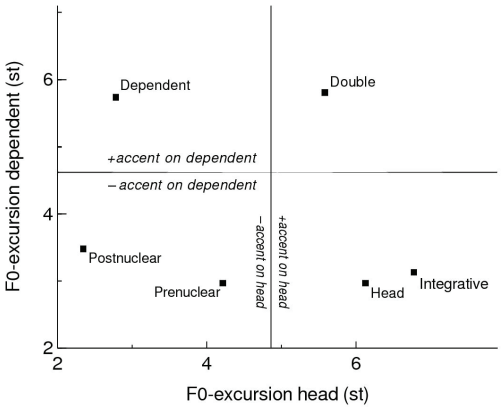


Figure 1. Excursion size of pitch movement on head (horizontally, semitones) and on dependent (vertically) for six focus distributions.

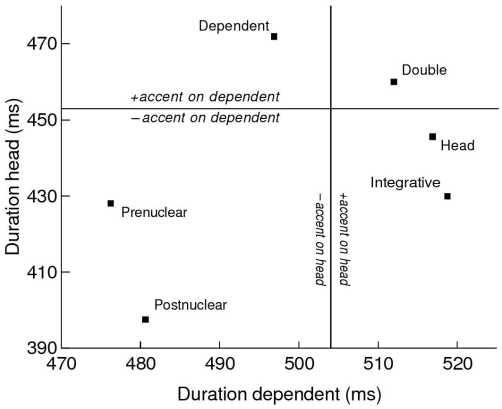


Figure 2.Duration of head (horizontally, ms) and on dependent (vertically) for six focus distributions.

It should be noted, finally, that the F_0-configuration on our target constituents is never completely flat, not even in [–focus, –accent] conditions; one reason for this is that there will always be some degree of downtrend, and, but more often than not, the speakers produce a small rise-fall, which would be the unintentional and involuntary response of the glottal mechanism to the aerodynamics of syllable production, or even a voluntary (but non-accent lending) movement, not exceeding 4 st.

Figure 2 presents the results for the duration (in ms) of the prosodic head and of the dependent, analogous to figure 1. The duration data correspond closely to the pitch data. Most importantly, there is no indication at all that an unaccented constituent that is part of an integrative focus, i.e., [+focus, –accent] is lengthened; it is, in fact, even shorter than when the head is in narrow focus. Moreover, time-expansion of the unaccented dependent in the integrative focus domain is absent, even though the dependent is often realised with a small (non-accent lending) pitch movement (between 2 and 5 st).

We conclude, therefore, that Eefting (1991) was basically correct, even if her materials were non-optimal. Her conclusion stands firm: there is no direct effect of plus versus minus focus on the duration of a prosodic constituent. Focus determines where the pitch accent goes, and only the accented word is lengthened, not the dependent constituent, even if it is part of the integrative focus around the head.

These data run counter to predictions derived from Lindblom's H&H theory. Focussed elements, even when they themselves are not accented, are communicatively important, and yet they are temporally reduced. Our speakers' behaviour does not reflect on-line estimation of the hearer's needs; rather, it seems that the predictions of the H&H theory have been fossilized in the phonetic implementation of a pitch accent. Since pitch accents very often (but not

always) occur on [+focus] words, and [+focus] words typically (but not always) bear pitch accents, a phonologized strategy to lengthen just the accented words, is generally sufficient to meet the hearer's communicative requirements.

References

Eefting, W.Z.F. (1991). 'The effect of information value and accentuation on the duration of Dutch words, syllables and segments.' *Journal of the Acoustical Society of America* 89, 412-414.

Fowler, C.A., Housum J. (1987). 'Talkers' signalling of "new" and "old" words in speech and listeners' perception and use of the distinction'. *Journal of Memory and Language* 26, 489-504.

Hart, J. 't, Collier, R., Cohen, A. (1990) *A perceptual study of intonation*. *An experimental-phonetic approach to speech perception*. Cambridge University Press, Cambridge.

Hermes, D.J. (1988) 'Measurement of pitch by subharmonic summation.' *Journal of the Acoustical Society of America* 83, 257-264.

Heuven, V.J. van (1994). 'What is the smallest prosodic domain?' In P. Keating (ed.) *Papers in Laboratory Phonology III: phonological structure and phonetic form*, Cambridge University Press, London, 76-98.

Hunnicut, S. (1985). 'Intelligibility versus redundancy – conditions of dependency.' *Language and Speech* 28, 45-56.

Lieberman, Ph. (1963). 'Some effects of semantic and grammatical context on the production and perception of speech.' *Language and Speech* 6, 172-187.

Lindblom, B.E.F. (1990). 'Explaining phonetic variation: a sketch of the H&H theory.' In W.J. Hardcastle, A. Marchal (eds.) *Speech production and speech modelling*, Kluwer, Dordrecht, 403-439.

Nooteboom, S.G. (1972). *Production and perception of vowel duration, a study of durational properties of vowels in Dutch*. Doctoral dissertation, Utrecht University.

Rietveld, A.C.M., Heuven, V.J. van (1997). *Algemene Fonetiek [General Phonetics]*. Coutinho, Bussum.

Rijnsoever, P.A. van (1989). 'From text to speech – user manual for diphone speech program DS.' *IPO Manual nr. 88*, Institute for Perception Research, Eindhoven.

Rooij, J.J. de (1979). *Speech punctuation. An acoustic and perceptual study of some aspects of speech prosody in Dutch*. Doctoral dissertation, Utrecht University.

Sluijter, A.M.C., Heuven, V.J. van (1995). 'Effects of focus distribution, pitch accent and lexical stress on the temporal organisation of syllables in Dutch.' *Phonetica* 52, 71-89.

Sluijter, A.M.C., Heuven, V.J. van (1996). 'Spectral balance as an acoustic correlate of linguistic stress.' *Journal of the Acoustical Society of America* 100, 2471-2485.

Son, R.J.J.H, Pols, L.C.W. (1990). 'Formant frequencies of Dutch in a text, read at fast and normal rate.' *Journal of the Acoustical Society of America* 88, 1683-1693.

This paper appeared in: Sjef Barbiers, Johan Rooryck, Jeroen van de Weijer (eds.) *Small words in the big picture. Squibs for Hans Bennis*. HIL Occasional Papers 2. Holland Institute of Generative Linguistics, Leiden, 1997, 37-42.