Intonation of ‘now’ in resolving scope ambiguity in English and Dutch

1,2Bettina Braun and 2Aoju Chen

1University of Konstanz, Germany
2Max Planck Institute for Psycholinguistics, Nijmegen, NL

Short Title: Intonation of now

Address for correspondence:
Bettina Braun
Department of Linguistics
University of Konstanz
Universitätsstr. 10
Fach 186
78467 Konstanz
bettina.braun@uni-konstanz.de
Tel: 0049-7531-882386
Abstract
The adverb *now* in English (*nu* in Dutch) can draw listeners’ attention to an upcoming contrast (e.g., ‘Put X in Y. *Now* put X in Z’). In Dutch, but not English, the position of this sequential adverb may disambiguate which constituent is contrasted. We investigated whether and how the intonational realization of *now/nu* is varied to signal different scopes and whether it interacts with word order. Three contrast conditions (contrast in object, location, or both) were produced by eight Dutch and eight English speakers. Results showed no consistent use of word order for scope disambiguation in Dutch. Importantly, independent of language an unaccented *now/nu* signaled a contrasting object while an accented *now/nu* signaled a contrast in the location. Since these intonational patterns were independent of word order, we interpreted the results in the framework of grammatical saliency: *now/nu* appears to be unmarked when the contrast lies in a salient constituent (the object) but marked with a prominent rise when a less salient constituent is contrasted (the location).

Keywords: Intonation, scope ambiguity, word order, information structure, contrast, Dutch, English
1 Introduction
Ambiguity is inherent in the use of language. A common type of ambiguity arises from the different ways in which words can be grouped into larger chunks. As a consequence, the same string of words can have different interpretations. For instance, the sentence *The teacher said John was intelligent* might convey that either John or the teacher was intelligent. In spoken language, prosody – in particular phrasing – provides a useful cue to resolve such ambiguities (e.g., Price, Ostendorf, Shatuck-Hufnagel & Fong, 1991; Snedeker & Trueswell, 2002; Speer, Kjelgaard & Dobroth, 1996). In the example above, a break after *said* conveys that John was intelligent, but a break before and after *said* conveys that the teacher was intelligent.

Ambiguities can also arise when the meaning of a certain word can apply to different parts in a sentence (i.e. scope ambiguity), as in the case of focus particles such as *even, only, and also*. The exhaustive particle *only*, for instance, can refer to *Bill* - Bill being the only person introduced to Mary (1a), to *Mary* - Bill being only introduced to one person, namely Mary (1b), or to both *Bill* and *Mary* - John only introduced two people to each other, namely Bill to Mary (1c). Prosody is again effective in resolving these ambiguities. Speakers can specify the intended interpretation of the utterance by assigning a falling accent to the constituent(s) that fall in the scope of *only* – marked by squared brackets in (1) (hereafter we refer to accent placement and accent type generally as intonation). Furthermore, 1a and 1b can be differentiated by means of phrasing, inserting an intonational phrase after *Bill* or not. The intonation of a particle itself can also play a role in scope marking, as in the case of German *auch* (‘also’) (e.g., Reis & Rosengren, 1997; Krifka, 1991).

(1)  a. John only introduced [Bill] to Mary.
    b. John only introduced Bill [to Mary].
    c. John only introduced [Bill] [to Mary].

Scope ambiguity induced by focus particles has received abundant attention in the literature, in terms of formalization (e.g., Altman, 1976; Bayer, 1996; Büring & Hartmann, 2001; König, 1994; Rooth 1992), processing (e.g., Crain, Ni & Conway,
1994), and acquisition (e.g., Dimroth & Watorek, 2000; Paterson, Liversedeg, Whilte, Filik & Jaz, 2006; Berger, Müller, Höhle, & Weissenborn, 2007). Much less is known about scope ambiguities caused by the adverb *now*.

Now may have three functions (e.g., Shiffrin, 1987; Hirschberg & Litman, 1993, Halliday & Hassan, 1976). It can be used as a temporal adverb, indicating that an event is taking place or should take place at the time when the utterance is produced as in *Let’s discuss that now*! (Bartsch, 1972). Further, it can be used as a discourse marker (or ‘cue phrase’ in Hirschberg & Litman’s terms) and signal the introduction of a new subtopic. Finally, *now* also signals a temporal sequence (e.g., *He already did A and now he does B*), and parallels the past tense sequential adverb *then* (e.g., *He first did A and then he did B*). Hirschberg and Litman (1993) showed that the first two functions can be differentiated intonationally. They analyzed productions from a radio call-in program and reported differences in phrasing and pitch accent realization for the temporal and discourse usage. In its temporal usage, *now* appeared mostly at the end of intermediate phrases and was usually accented. In its discourse usage, *now* was mostly positioned phrase-initially and left unaccented.

In this study, we focus on the usage of *now* as a sequential adverb. More specifically, we investigate how speakers vary the prosodic realization of *now* to signal the scope of the upcoming contrasts in game instructions as in (2). These kinds of instructions (especially 2b and 2c) are frequently used in psycholinguistic experiments investigating the role of accent placement and accent type in online reference resolution (e.g. Dahan, Tanenhaus & Chambers, 2002, Chen, den Os & de Ruiter, 2007).

(2) a. Put the zebra in cell 1; now put [the baby] in cell 1. (OBJECTCONTRAST)
b. Put the zebra in cell 1; now put the zebra [in cell 2]. (LOCATIONCONTRAST)
c. Put the zebra in cell 1; now put [the baby] [in cell 2]. (DOUBLECOnTRAST)

Irrespective of the locus and number of contrasts, *now* needs to be placed phrase-initially in English. Therefore, English speakers are likely to vary prosody (pitch accent type and phrasing) of *now* to disambiguate the locus and number of contrasts. Indeed, an initial indication for such uses of intonation can be found in the experimental materials of
the above-mentioned psycholinguistic studies. In Dahan et al.’s (2002) American English materials, *now* was spoken with an H* pitch accent when the object was contrastive (2c) but mostly with a rising accent (either L*+H or L+H*) when the object was maintained (2b). In the study by Chen et al. (2007) on British English, *now* was mostly unaccented or realized with a weak rise when the object was contrastive (2c) but with a strong rise when the object was maintained (2b). These production data (though produced only by a single speaker each) suggest that the intonation of *now* varies depending on the information status of the upcoming object. However, in the stimuli of both studies, a contrastive object was always moved to a new location (resulting in a double contrast as in example 2c). Therefore, it remains to be tested whether the observed patterns hold for more than one speaker and whether speakers also vary the prosodic realization of *now* to differentiate a contrast in the object from a double contrast.

Secondly, this study investigates whether underlying word order flexibility affects the use of prosody in this regard. Previous studies on the linguistic markings of information structure have shown that speakers not only employ prosody but also word order and particles (Keller & Alexopoulou, 2001 for Greek; Kim & Avelino 2003 for Mexican Spanish, Göbbel 2003 for Romanian, Zubizarreta 1998 for Germanic and Romance languages in general, Vallduví & Engdahl, 1996 for German, Dutch, English, Catalan, Hungarian and Turkish, Swerts & van Wijk, 2005 for Dutch). Therefore, prosody might not be used at all (or to a lesser extent) if a language allowed speakers to disambiguate the contrast conditions by means of word order. A language par-excellence to test this assumption is Dutch. Dutch has a very similar intonational grammar as English (Gussenhoven, 1988; ’t Hart et al., 1990) and the same intonational flexibility (e.g., Vallduví & Engdahl, 1996; Ladd 1996). More importantly, informal inquiries suggest that at least the locus of the contrast (OBJECT-CONTRAST vs. LOCATION-CONTRAST) can be disambiguated by means of word order: *nu* (‘now’) can optionally be placed prior to the contrasted constituent (as shown in example (3)).

(3) a. Verplaats de zebra naar vak 1; verplaats *nu* [de baby] naar vak 1.  
     lit: *Put the zebra in cell 1; put now the baby in cell 1.*  

b. Verplaats de zebra naar vak 1; verplaats de zebra *nu* [naar vak 2].
2. Word order preferences in English and Dutch

English and Dutch participants were presented with instruction pairs similar to those in example (3) and rated the acceptability of the second instruction.

2.1 Methods

2.1.1 Materials

Eighteen pairs of instructions were constructed for Dutch and English each. The position of the adverb in the second instruction (phrase-initial, after the verb, and after the object) and contrast condition (LOCATION CONTRAST, OBJECT CONTRAST, and DOUBLE CONTRAST) were varied orthogonally. Object nouns were highly frequent, every-day words. Half of them were monosyllabic, half disyllabic.

2.1.2 Participants

Twenty-eight native English and thirty native Dutch speakers participated in the experiment. English participants were made aware of the web experiment through personal networks in London and Chicago. Nineteen were from the US and 9 from the UK. The Dutch participants were recruited via the subject pool from the Max Planck Institute for Psycholinguistics in Nijmegen (MPI). All participants (28 female, 30 male) were between 19 and 46 years old. Participants could freely decide where to perform the web experiment.
2.1.3 Procedure
Two lists were constructed; the presentation of the stimuli was pseudo-randomized. Two subsequent trials differed in at least one factor (adverb position or contrast condition). Participants received instructions that the experimental sentences were written by a non-native speaker for a board game in which objects had to be moved to different positions. They were told that the non-native speaker was not allowed to use pronominalized forms and that their task was to rate the acceptability of the second instruction on a scale from 1 (not acceptable) to 5 (highly acceptable). After they had entered their rating via the keyboard of their PC, the next sentence pair appeared. The experiment lasted approximately five minutes.

2.2 Results
One Dutch and one English participant who took less than less than 90 seconds overall (less than 5 seconds per trial) and one Dutch participants who took more than 15 minutes overall were excluded as outliers. Ratings were analyzed using multi-level regression models with adverb position, contrast condition, and language as fixed factors and participant as random factor (Baayen, 2008). Data points for which the absolute standardized residuals were greater than 2.5 were removed as outliers and the model was refitted. P-values were estimated as posterior probabilities based on a Markov Chain Monte Carlo simulation with 10,000 runs.

Results showed a three-way interaction between adverb position, contrast condition, and language (p < 0.0001). In what follows we first present the results from English participants, followed by those from Dutch participants.

For English participants, there were main effects of contrast condition (p < 0.05) and adverb position (p < 0.0001), but no interaction (p > 0.1). To ensure that the interaction term does not improve the fit of the model, we compared the predictive power of the full model with the one without the interaction term. The log-likelihood ratio for the simpler model was 566.1, compared to 563.2 for the model with the interaction (χ²(4) = 5.9, p > 0.2). This supports the interpretation that there is no interaction between the factors. Ratings for trials with a LocationContrast were rated less acceptable than trials with a DoubleContrast (β = -0.2, Lower Bound = -0.36, Upper Bound = -0.04, p < 0.05). More importantly, positioning the adverb phrase-initially was perceived as more
acceptable (mean rating 4.7) than positioning the adverb after the object (mean rating 3.0, p < 0.0001), which in turn was perceived as more acceptable than placing it after the verb (mean rating 2.4, p < 0.0001), see Figure 1.

Please insert Figure 1 about there.

For Dutch participants, there were main effects of adverb position (p < 0.0001), contrast condition (p < 0.0001), and an interaction between the two (p < 0.0001), see Figure 2. Overall, nu in phrase-initial position was considered worst, independent of contrast condition (mean rating 2.1 compared to 4.2 for positioning the adverb after the verb, p < 0.0001 and 4.0 for positioning it after the object, p < 0.0001). Dutch listeners judged sentences as most acceptable if nu was positioned after the verb but more so when there was a contrast in the location or a double contrast and less so when there was a contrast in the object. Placing nu after the object received highest ratings in OBJECTCONTRAST (mean rating 4.4, p < 0.05) and lowest in LOCATIONCONTRAST (mean rating 3.1, p < 0.01).

Please insert Figure 2 about here

2.3 Discussion
Rating data from the two web experiments showed that English is indeed more restricted in word order than Dutch when it comes to positioning the adverb now/nu (hereafter now when we refer to it language-independently). English participants preferred sentences in which now was placed sentence-initially, independent of contrast condition. Dutch participants’ preference, on the other hand, was dependent on contrast condition. In sentences with a contrast in location, they preferred nu after the verb, but in OBJECTCONTRAST and DOUBLECONTRAST sentences, they judged nu after the object as more acceptable. The fact that these results differ from our initial impressions shows that there is more than one way to mark upcoming contrasts in Dutch. Possibly, participants’ preference is dependent on the prosodic structure projected during reading (e.g., Fodor,
and on the willingness to tolerate unexpected productions by a non-native speaker of Dutch. But importantly, these experiments have provided evidence that English and Dutch differ with respect to constraints on word order for the adverb NOW.

3 English Production Experiment

A movie clip description task (cf., Terken & Hirschberg, 1994; Speer, Warren, & Schafer, 2003, Swerts, 2007) was used to elicit sentences like the examples in (2). Participants watched 15 movie clips depicting two objects on a 3 x 3 grid that were moved to different locations. Their task was to describe the action they saw in the form of an instruction to someone else after the action was completed.

3.1 Methods

3.1.1 Materials

Twelve pairs of familiar objects were selected. For every object pair, three movie clips were constructed, resulting in 36 movie clips in total. Each movie clip consisted of two parts. Initially, the screen depicted a green 3 x 3 grid (1345 x 1030 pixels) with nine cells. Each cell contained the respective cell number in black Arial 12pt in the top left corner, highlighted with pink background colour. The grid contained two objects (white line drawings on black background, 96 x 96 pixels each) which were placed together in one of the cells (a black-white drawing of the display is shown in Figure 3). In the first part of the movie clip, one of the two objects was moved to another cell. For the second part of the movie clip there were three possibilities. In OBJECTCONTRAST, the other object was moved to the same location as the first object was moved to. In LOCATIONCONTRAST, the same object was moved again but to another location. In DOUBLECONTRAST, the other object was moved to another empty cell. For each object pair, the object moved in the second part of the movie clips was identical across conditions.

The movie clips were created using FinalCut Pro and exported as a self-contained movie to a QuickTime file (DV PAL 48 kHz). The first movement started 200 ms after the display of the grid containing the objects and took 3 seconds. This was followed by a
pause of four seconds in which participants could phrase the first instruction. The second
movement also took 3 seconds and was followed by a 4-second pause so that participants
had time to phrase the second instruction.

3.1.2 Participants
Eighteen Southern British English speakers (9 female and 9 male) took part in the
experiment. Sixteen of them were students at the University of Sussex; two were
exchange students who just arrived in Nijmegen at the time of testing. Participants were
unaware of the purpose of the experiment and received a small fee for their participation.

3.1.3 Procedure
Participants were seated in front of a computer screen. They were informed that they
were about to see a number of movie clips consisting of two parts and that they should
describe the actions they saw in the form of an instruction to another person so that
he/she could perform the respective actions. Therefore, participants were asked to speak
clearly and at a normal speaking rate. In addition, they were instructed to refer to the
object by its name (the first word that came to their mind) and to refer to the cell by using
the number printed in the cell (1 to 9). Participants were also told to use the word put for
describing the action, the word now to mark the second instruction, and to avoid the word
also.

Three experimental lists were compiled, each containing 12 movie clips, four in
each condition. The three lists together contained all the object pairs in all conditions. To
familiarize participants with the task, each list started with three practice movie clips, one
in each condition. In this practice session, participants were corrected if they did not
follow the experimenter’s instructions.

The experiment was conducted by means of the NESU (Nijmegen Experiment
Setup) experimental software. Sixteen participants were tested individually in a quiet
room at the University of Sussex. The remaining two were tested in a sound-attenuated
booth at the MPI. Twelve pairs of instructions were elicited from each participant. The
productions were recorded on DAT-tapes with a sampling rate of 48 kHz and a resolution
of 16-bit.
3.2 Analysis
The second sentence in each pair of instructions was marked on the word level for the verb (put), the adverb (now), the target object to be moved, and the location (cell n or just the number). In what follows, we will refer to the instruction containing the adverb as ‘second instruction’. The second instructions were all produced in the order Now put the X into cell Y.

To ease subsequent prosodic analysis, we focused on the productions from the female speakers.1 One speaker had to be excluded because of poor recording quality, leaving 8 speakers for the analysis. Five utterances from these speakers had to be discarded because of hesitations and absence of the adverb. Following this selection procedure, 89 second instructions were analyzed further. There were 32 utterances in LOCATIONCONTRAST, 29 in OBJECTCONTRAST, and 28 in DOUBLECONTRAST.

To ensure that speakers appropriately encoded the intended information structures, we subsequently recruited five native speakers of English (students at the University of York) to evaluate the information structure in the 89 second instructions. These judges indicated for each of these second instructions whether they expected the preceding first instruction, which was not presented to them, to differ in the object, the location, or both. The presentation of utterances was blocked by speaker and the utterances were randomized within a speaker. Listeners were blind to contrast condition. Utterances were classified correctly by the majority of listeners (three or more) in 85.9% of the cases. The small number of misclassifications was restricted to interpreting the intended object or location contrast as a double contrast.

3.2.1 Phonological analysis
The 89 second instructions were prosodically annotated following IViE (Grabe, 2004), a notation system proposed for varieties of British English and originated from ToDI (Transcription of Dutch Intonation, see Gussenhoven, 2005). In IViE, only differences in the tonal movement on and following the stressed syllable (i.e. ‘trailing tones’) and the tonal movement at intonational phrase boundaries (corresponding to break index 4 in the ToBI system) are transcribed. The sentences were first annotated by the authors

---

1 Impressionistic analysis suggests similar patterns for the male participants. But their production showed smaller f0-range and was therefore harder to analyse than the female participants’ productions.
independently and blind to contrast condition. Then, the labels were compared and
doubtful cases as well as disagreements (< 10% of the cases) were resolved by the
authors together. Three accent types occurred in *now*, (H* - high level, H*L - fall, and
L*H - low rise, for a schematic representation of these accent types see Table A1 in the
Appendix).

Example pitch tracks of English utterances are shown in Figures 4 and 5.

Please insert Figure 4 about here.

Please insert Figure 5 about here.

3.2.2 Phonetic analysis
To further corroborate the phonological analysis of *now* – which was relatively difficult
due to its shortness and the subsequent compression in the realization of the tonal targets
(Grabe, 1998) – we analyzed the realizations of *now* phonetically. Since *now* is fully
sonorant, the fundamental frequency contour of the adverb was extracted automatically at
5ms steps using Praat (Boersma & Weenink, 2009). The first five f0-values (from the
first 20 ms in *now*) were not considered because of potential f0-perturbations. The f0-
maximum in the adverb was identified automatically. The f0-exursion of the rise was
calculated from the first available f0-value of the adverb to the f0-maximum in the adverb
and was converted to semitones. The f0-exursion of the fall was calculated from the f0-
maximum in the adverb to the f0-value at the mid point of the following syllable (‘put’)
and converted to semitones. The mid piont of the following syllable was used because the
adverb itself was too short to determine a fall and since the mid point of the following
syllable could be reliably extracted. Furthermore, the alignment of the f0-maximum with
respect to the onset of *now* was measured. Finally, we calculated the relative duration of
the adverb (duration of the adverb divided by the duration of the verb). This
normalization was done to minimise effects due to different speech rates. The complete
set of alignment, scaling and duration information is summarized in Table A2 in the Appendix.

All these acoustic variables were subjected to individual multi-level regression models with accent type as fixed factor (H*, H*L, L*H), and speaker and target words as crossed random variables. Since the variance for target words was effectively zero, this random variable was dropped. Data points for which the absolute standardized residuals were greater than 2.5 were removed as outliers and the model was refitted. P-values were estimated as posterior probabilities based on a Markov Chain Monte Carlo simulation with 10000 runs. If the factor pitch accent type was significant, separate posthoc comparisons between accent types were conducted by resetting the reference level in the model (p-values were corrected for multiple analyses using Bonferroni adjustment). Here we present only the minimal set of acoustic differences that are necessary to distinguish between all accent types. The f0-excursion of the rise differentiated the accents H*, H*L, and L*H (see Figure 6, corrected p-value 0.025), but not H* and no accent. However, an adverb with an H* accent differed from an unaccented adverb in its relative duration (1.22% of the duration of the verb for an accented now compared to 1.03% for an unaccented now, p < 0.05).

Please insert Figure 6 around here

3.3 Results

As the adverb now and the object noun were separated by only two unstressed syllables, they might influence each other in intonational realization. We therefore consider the intonation realization of the object noun first and then the intonational realization of now. Finally, we will present a matrix showing the co-occurrences of pitch accent types in now and the object noun.

*Intonational realization of the object noun.* The distribution of intonation patterns in the object noun across conditions is presented in Table 1. As expected, the noun was mostly accented with H*L when the object was contrasted (89.7% in ObjectContrast and 82.1% of the cases in DoubleContrast). When the object was not contrasted, it was mostly left unaccented (59.4% of the cases) or was accented with L* (21.9% of the cases). As...
regards intonational phrasing, participants did not vary phrasing to signal the different contrast conditions. They produced a break in 10 out of 29 cases in OBJECTCONTRAST, in 7 out of 32 cases in LocationContrast, and in 5 out of 28 cases in DOUBLECONTRAST.

Please insert Table 1 around here.

*Intonational realization of the discourse marker *now*. There were only four instances of an intonational phrase break following the adverb, two in LOCATIONCONTRAST and two in OBJECTCONTRAST. The absolute total number of tokens of each intonation pattern of *now* is provided in Table 2. As the total number of trials differed across conditions and across speakers, and speakers did not use all the intonation patterns equally frequently, we also report the mean percentage (%) distribution of each intonation pattern in each condition in Table 2. The mean % distribution of a given intonation pattern in a given condition was computed by averaging the percentages of the trials spoken with that accent type in the respective condition from all speakers. The most common patterns were 'no accent' and H*. In LOCATIONCONTRAST, L*H, H*, and 'no accent' occurred with similar frequency. In OBJECTCONTRAST and DOUBLECONTRAST, *now* was most frequently unaccented.

Please insert Table 2 around here

To verify the distribution of intonation patterns in *now* over contrast conditions statistically, we carried out a multinomial logistic regression analysis (hereafter MLR) at the significance level of 0.05. The independent variable (or the predictor variable) was *contrast condition* with three categories, LOCATIONCONTRAST, OBJECTCONTRAST, and DOUBLECONTRAST. The dependent variable (or the outcome variable) was the intonation of *now*, consisting of four categories (H*, H*L, L*H, and 'no accent'). The accent L*H was used as the reference category because the distribution of L*H appeared to vary more clearly with contrast conditions compared to the other accent types.²

² Note that it was not possible to include speaker as a predictor variable because not every accent type was produced by every speaker.
The by-chance accuracy in assigning the intonation patterns to the contrast conditions (calculated by summing up the squared proportion of cases of each intonation pattern in the data) was 39.8%. The model fitting was improved significantly after the variable contrast condition was added (fit: 47.2%, $\chi^2 = 14.60$, df = 6, p < 0.05). This improvement supports a significant overall relationship between the contrast condition and the intonation of now.

The Wald statistics, comparable to the t-statistics in linear regressions, revealed that contrast condition significantly predicted the choice between an H* and an L*H accent on now in ObjectContrast ($b = 2.30$, Wald = 4.01, df = 1, p < 0.05) compared to LocationContrast, and the choice between 'no accent' and L*H in ObjectContrast ($b = 2.25$, Wald = 6.20, df = 1, p < 0.05) and DoubleContrast ($b = 2.81$, Wald = 6.60, df = 1, p < 0.01) compared to LocationContrast. The odds ratios (Exp(B) values in the SPSS output) indicated first that the odds of now being spoken with H* compared to the odds of now being spoken with L*H was 10 times higher in ObjectContrast than in LocationContrast. Second, the odds of now being spoken with 'no accent' was 16.67 times higher than the odds of now being spoken with L*H in ObjectContrast than in LocationContrast. Third, the odds of now being spoken with 'no accent' was 9.44 times higher than the odds of now being spoken with L*H in DoubleContrast than in LocationContrast.

---

5 The absence of L*H in the object condition made it impossible to examine the role of contrast condition in predicting the use of no accent in now compared to the use of L*H. We therefore introduced one fictive instance of now, accented with L*H, into the object contrast condition by replacing one instance of unaccented now with L*H and ran the MLR analysis. This operation caused little changes to the Wald statistics concerning H* and H*L but generated interpretable results for the category no accent.

6 Odds are defined as the probability of an event occurring divided by the probability of an event not occurring. The odds ratio is the proportionate change in odds, calculated by dividing the odds after a unit change in the predictor by the odds before that change. It serves as an indicator of the change in odds resulting from a unit change in the predictor, similar to the $b$ coefficient but is easier to interpret because it does not involve a logarithmic transformation. If the odds ratio is larger than 1, it indicates that as the predictor increases, the odds of the outcome occurring increases. If the odds ratio is smaller than 1, it indicates that as the predictor increases, the odds of the outcome occurring decreases (Field, 2009).

6 Conversely, the odds of now being spoken with L*H compared to the odds of now being spoken with H* was higher in LocationContrast. This was somewhat unexpected given that H* and L*H seemed to occur with similar frequency (see Table 1). A closer inspection of the within-speaker distribution of L*H and H* revealed that only four speakers produced L*H and five speakers produced H*. Importantly, L*H was the most frequent accent type in LocationContrast in all the four speakers who produced this accent, whereas H* was the most frequent accent type in only two out of the five speakers who used this accent. This suggests that the preference for L*H over H* in LocationContrast was a speaker-specific effect.
Co-occurrence of pitch accents. Table 3 shows the frequency with which a certain pitch accent on *now* is followed by a certain pitch accent in the object noun to address the issue of whether the observed intonation patterns in *now* might be due to the avoidance of an accent clash.

Please insert Table 3 about here.

As can be seen in Table 3, an accented *now* is followed by an accented object noun in 34 cases but by an unaccented object noun in only 14 cases. Conversely, an accented object noun was preceded by an accented *now* in 33 cases and by and unaccented *now* in 36 cases.

3.4 Discussion
These results show that speakers varied the intonation of *now*, but not phrasing, to disambiguate the locus of the contrast (i.e., whether the object was contrasted or not). The number of contrasts, however, was not signalled at all. By and large, *now* was preferably unaccented or accented with H* when the object was contrasted, independent of whether the location was contrasted at the same time. On the other hand, *now* was preferably accented with L*H when the object was not contrasted. Intonation of *now* hence only signalled whether the object was contrasted or not, similar to the patterns we observed in the experimental materials used by Dahan et al. (2002) and Chen et al. (2007).

Strikingly, English speakers did not use phrasing or the intonation in the object noun to differentiate between a contrast in the object and a double contrast in the object and the location. In both cases, the object noun was realized with a pitch fall (H*L), sometimes with a phrase break, but mostly without. Therefore it is not surprising, that native speakers of English sometimes had difficulties telling these two conditions apart.

Since a contrasted object is mostly accented, the preference for an unaccented *now* in conditions with a contrasted object might be phonologically conditioned: To avoid an accent clash between *now* and the object noun, the realization of *now* may be toned
The opposite can be argued for the frequent use of a prominent rise (L*H) in *now* in LOCATION\textsc{Contrast}, where the object noun was mostly unaccented or accented with a L*, which made it possible for *now* to be accented. However, this accent-clash-avoidance assumption is not born out by the data. As can be seen in the co-occurrence matrix of pitch accents in *now* and the object in Table 3, an accented *now* was more often followed by an accented object noun than by an unaccented one. Conversely, an accented object was equally often preceded by an accented *now* as by an unaccented *now*.

Hence, an unaccented *now* appears to be a marker for an immediate contrast (in the present materials a contrasting object), while an L*H signals that the contrast will appear only later in the sentence (in our materials a contrast in location). Such a function of L*H in *now* fits well with the function of a rising nuclear contour (a rising nuclear pitch accent plus a high boundary tone) as a cue to continuation in general in English (Cruttenden, 1997, Pierrehumbert, 1981). Alternatively, since the salience of these two constituents, the object and the location, is not the same (in syntactic theory, objects are usually treated as more salient than local adverbials, cf. obliqueness hierarchy by Keenan & Comrie 1977), an unaccented *now* can be interpreted as a marker for a contrast in a salient constituent (in the present materials the object), and L*H as a marker for a contrast in a less salient constituent (e.g., the location). However, the present data do not allow us to decide between these two accounts as sequential order is confounded with grammatical role (i.e., the object always preceded the location).

---

7 Note that in IViE, H* is a high level tone or a slightly rising contour, which is perceptually a relatively weak accent.
4 Dutch production experiment

4.1 Methods

4.1.1 Materials
The Materials were identical to the ones described in 3.1.1

4.1.2 Participants
Twenty-two native Dutch speakers (11 female and 11 male, students at Radboud University Nijmegen) took part in the study. They were unaware of the purpose of the experiment and received a small fee for their participation.

4.1.3 Procedure
The procedure was identical to the one described in 3.1.3. Dutch participants were tested in a sound-attenuated booth at the MPI in Nijmegen. The productions were recorded on DAT-tapes with a sampling rate of 48 kHz and a resolution of 16-bit.

4.2 Analysis and Results

4.2.1 Word Order
The second sentence in each pair of instructions was annotated on the word level for the verb (Verplaats), the adverb (nu), the target object to be moved, and the location (vak n or just the number). Two word orders occurred in Dutch speakers’ productions, Verplaats nu de X naar vak Y (lit: ‘Put now the X in cell Y’) and Verplaats de X nu naar vak Y (lit: ‘Put the X now in cell Y’). The former was the dominant one across conditions and was used in 89% of the cases. At first, this result looked striking, given that Dutch speakers have the option to exploit word order variations to partly disambiguate the conditions. At a second glance, however, it was not completely unexpected as the dominant word order was the one that received the highest acceptability scores overall. Some speakers used one word order throughout, independent of contrast condition (3 males and 9 females, ‘fixed word order group’ henceforth), others made use of word order variation (8 males and 2 females, ‘variable word order group’ henceforth).

We will first discuss the intonational realization of nu for the ‘fixed word order group’ in the same fashion as done for English now to answer the question whether
underlying differences in word order flexibility affect the degree to which prosody is used to disambiguate the contrast conditions. Then we will present a descriptive intonational analysis of the productions from the ‘variable word order group’ to get an insight of how choice of word order interacts with the prosodic realization of \textit{nu}.

4.2.2 ‘Fixed word order group’ (Verb \textit{nu} Object Location)

To have a sample size comparable to the data analyzed for English, we selected 8 female speakers who only produced the dominant word order. The Dutch productions analyzed here hence only differed from the English ones in the relative order between verb and adverb.

Sixteen of the second instructions had to be discarded because of hesitations and absence of the adverb. Following this selection procedure, 80 second instructions were selected for further analysis (28 in \textit{LOCATION\textsc{Contrast}}, 24 in \textit{OBJECT\textsc{Contrast}}, and 28 in \textit{DOUBLE\textsc{Contrast}}). Similar to the English data, these second instructions were presented to five native speakers of Dutch for evaluation. The evaluation procedure was the same as described for English. The utterances were classified correctly by the majority of these judges (three or more) in 83.8\% of the cases, which is similar to the classification rate for English utterances (85.9\% of the cases). The small number of misclassifications was again restricted to interpreting the intended object or location contrast as a double contrast.

4.2.2.1 Phonological analysis

The 80 second instructions were prosodically annotated following ToDI. Like its English counterpart IViE described in 3.2.1, ToDI only transcribes the tonal movement on and following the stressed syllable and the tonal movement at intonation phrase boundaries. However, initial inspection of the pitch tracks revealed differences in the tonal movement leading up to the starred tone of \textit{nu}. These differences were caused by the tonal transition from the verb preceding the adverb to the adverb itself (\textit{verplaats \textit{nu}}) and might therefore not be captured by the phrase-initial boundary tone (as it does not sufficiently specify the tonal realization of the verb). In order not to lose potentially useful information for the
intonation of *nu*, we decided to extend the ToDI categories by adding two more accent types (LH* - a high level tone preceded by a low tone, and LH*L - a fall preceded by a low tone). There were hence five accent types for the intonation of *nu* (see Table A1 in the Appendix). The annotation procedure was the same as described for English.

Example pitch tracks are shown in Figures 7 and 8.

4.2.2.2 Phonetic analysis

The f0-excursion of the rise was calculated from the mid of the stressed vowel /a/ of the verb *verplaatst* to the f0-maximum in *nu*. The other variables were identical to the English measurements. The predictors and the model fitting were identical to that of the English data. Figure 9 shows box plots of the f0-excursion of the rise for all accent types, including ‘no accent’. For this analysis, posthoc comparisons were conducted with H*, LH*, and L*H as reference levels. The corresponding p-value was corrected to 0.016. Results showed that f0-excursion of the rise was largest for L*H, followed by LH* and LH*L, and smallest for ‘no accent’, H* and H*L.

As can be seen in Figure 10, the two pairs of accent types that did not differ in the f0-excursion of the rise (LH* and LH*L as well as H* and H*L) differed in the f0-excursion of the fall. The excursion of the f0-fall was significantly larger in H*L compared to H* and in LH*L compared to LH*, i.e. the fall was significantly more pronounced in accent types with a low trailing tone (corrected p-value 0.025).

Final analysis showed that an adverb with an H* accent was significantly longer than an unaccented one (31% of the duration of the verb for an accented adverb compared to 26% of the duration of the verb for an unaccented one). A summary of the alignment, scaling, and duration measures is provided in Table A3 in the Appendix.

4.2.2.3 Results

Please insert Figure 9 around here

Please insert Figure 10 around here
The presentation of Dutch results follows the same ordering as for the English results.

**Intonational realization of the object noun.** The distribution of pitch accent types in the object noun across conditions is presented in Table 4. Similar to English, the most frequent accent type for ObjectContrast was H*L (70.8% of the cases). In DoubleContrast, participants used L*H most frequently (35.7% of the cases), followed by H*L (25% of the cases). For LocationContrast, the most frequent accent types were L*H (25.0% of the cases), followed by H* (21.4%) and 'no accent' (17.9%). Further, Dutch speakers, like English speakers did not vary phrasing after the object noun to differentiate the contrast conditions. They produced 15 breaks in LocationContrast and ObjectContrast (out of 28 and 24, respectively), and 12 breaks (out of 28) in DoubleContrast.

Please insert Table 4 around here.

**Intonational realization of the adverb *nu.** There were five instances of an intonational break following the adverb, two in LocationContrast and DoubleContrast each, one in ObjectContrast. The frequency of intonational breaks after *nu* was not affected by contrast condition ($\chi^2 = 0.3$, df = 2, p > 0.8). In what follows, we will present the results on pitch accent types of *nu*.

The intonational realizations of *nu* (absolute numbers and mean % distribution) are shown in Table 5. In LocationContrast, *nu* was mostly accented; the most common accent types were rising accents, i.e. LH* and L*H. In ObjectContrast, however, *nu* was most frequently unaccented. In the DoubleContrast condition, unaccented *nu* and *nu* accented with LH* occurred more frequently than the other patterns.

Please insert Table 5 around here.

As in the English data, an MLR analysis was conducted on the Dutch data at the significance level of 0.05. The dependent variable was the intonation of *nu*, consisting of six categories (H*, H*L, L*H, LH*, LH*L, and 'no accent'); like in English the reference
category was also set to L*H. The independent variable was contrast with three categories, LOCATIONCONTRAST, OBJECTCONTRAST, and DOUBLECONTRAST.

The by-chance accuracy in assigning the intonation patterns to the contrast conditions was 23.4%. The model fitting was improved significantly after the variable contrast condition was added (fit: 36.2%, $\chi^2 = 23.10$, df = 10, p = 0.01), indicating a significant overall relationship between the contrast condition and the intonation of nu.

The Wald statistics showed that contrast condition significantly distinguished the tokens of nu spoken with 'no accent' from those spoken with L*H (reference category) ($b = 4.34$, Wald = 8.45, df = 1, p < 0.005 in OBJECTCONTRAST, $b = 2.79$, Wald = 4.82, df = 1, p = 0.05 in DOUBLECONTRAST). The odds ratios showed that the odds of nu being spoken with 'no accent' was 77 times higher than the odds of nu being spoken with L*H in OBJECTCONTRAST compared to LOCATIONCONTRAST. Similarly, the odds of nu being spoken with 'no accent' compared to the odds of nu being spoken with L*H was 16 times higher in DOUBLECONTRAST than in LOCATIONCONTRAST. These results largely support the observations from the mean % distribution of accent types.

Co-occurrence of pitch accents. Table 6 shows the co-occurrence matrix between the intonational realization of the discourse marker and the pitch accent types observed on the object noun.

The table of co-occurrence is not provided, but it shows the relationship between accent types and their occurrence in different sentence structures.

Please insert Table 6 around here.

An accented nu was followed by an accented object noun in 54 cases but by an unaccented object noun in only 5 cases. Conversely, an accented object noun was preceded by an accented nu in 54 cases but by an unaccented nu in only 17 cases.

4.2.3 ‘Variable word order group’ (Verb Object nu Location and Verb nu Object Location)

Excluding 10 cases with hesitations or no nu, the 10 speakers from the ‘variable word order group’ provided 29 second instructions in the non-dominant word order, in which nu followed the object (e.g., ‘Verplaats de baby nu naar vak 1’), and 89 second instructions in the dominant word order (e.g., ‘Verplaats nu de baby naar vak 1’). One
female speaker produced the non-dominant word order throughout, independent of contrast condition. The other nine speakers used this word order more frequently in LOCATION CONTRAST (7 out of 31 cases) and DOUBLE CONTRAST (9 out of 31 cases) than in OBJECT CONTRAST (1 out of 27 cases), $\chi^2 = 7.4$, df = 2, p < 0.05.

The second instructions were prosodically annotated as described in section 4.2.2.1 The distribution of intonational patterns of $nu$ in the 89 trials with the dominant word order were similar to that reported for the trials produced by the fixed word order group in 4.2.2.3, see Table A4 in the Appendix.

In the 21 second instructions produced in the non-dominant word order, the speakers varied phrasing more across conditions. In LOCATION CONTRAST, there was mostly a break after the adverb, i.e. before the location (9 out of 11 cases), in OBJECT CONTRAST, there was mostly a break before the adverb, i.e. after the object (3 out of 5 cases), but there was no preference between these two options in DOUBLE CONTRAST (see Table A5 in the Appendix).

Regarding the intonational realization of the adverb in sentences with the non-dominant word order, we found that $nu$ was consistently produced with rising accents (mostly L*H but also H* followed by a high boundary tone) in LOCATION CONTRAST but was mostly unaccented in OBJECT CONTRAST. In DOUBLE CONTRAST there was no clear preference for one kind of accent over the others (see Table 7). These patterns were similar to the patterns in the sentences with the dominant word order produced by the fixed word order group.

Please insert Table 7 about here

4.3 Discussion

A third of the Dutch speakers appeared to make use of word order variation to signal the contrast condition. There were more instances in which $nu$ was placed before the location in LOCATION CONTRAST and DOUBLE CONTRAST than in OBJECT CONTRAST. The majority of the speakers placed $nu$ after the verb, independent of contrast condition. The preponderance of placing $nu$ after the verb is somewhat surprising, given the option to vary the word order to signal the different contrast conditions. Possibly, the repeated
production of instructions triggered syntactic priming (e.g., Bock, 1986), i.e., the preference to repeat the same word order across trials.

Independent of choice of word order, Dutch speakers varied the intonation in the object noun to differentiate a single contrast in the object from a double contrast. In double contrast, they used a higher proportion of rising accents (L*H), while they overwhelmingly used a pitch fall (H*L) in ObjectContrast. As argued in section 3.4, the use of the pitch rise may signal continuation, which may in this case be used as a hint for the listener that there is another contrast to come.

The use of phrasing covaried with the choice of word order. In the dominant word order (e.g., Verplaats nu het boek naar vak 3), Dutch speakers did not vary phrasing after nu to mark the contrast conditions. However, in the non-dominant word order (e.g., Verplaats het boek nu naar vak 3), they produced an intonational break more frequently after nu (and before the location) when the location was contrasted but more frequently before nu (and after the object) when only the object was contrasted. This result suggests that Dutch speakers had the tendency to prosodically isolate the contrasted constituent. Furthermore, it suggests that phrasing was used in tandem with word order scrambling to disambiguate contrast conditions – at least by some Dutch speakers. The finding that speakers not only altered word order but also prosody for information structural purposes is not uncommon. For example, in many languages, the fronting of focal elements (e.g., It was PETER who left) is accompanied by prosodic salience (e.g., Prince, 1978).

Interestingly, the use of intonation variation in nu to signal the locus of the contrast was largely independent of word order, similar to the use of intonational variation in the object noun. Nu was predominantly produced with rising accents (mostly L*H) when the location was contrasted; but it was mostly unaccented or spoken with H* when only the object was contrasted. Apparently, both word orders can be used in each of the three contrast conditions if the prosodic realization of nu (and the object noun) is chosen appropriately. This flexibility in word order may be the key to understanding the perplexing fact that the appropriateness ratings obtained from the web experiment and the distribution of the dominant and non-dominant word orders in actual productions do not resemble each other more. Positioning nu before the location is more likely to signal a contrast in location or a double contrast but at the same time, this word order (though
with a different prosodic realization) may also signal a contrasting object. Since prosody was not controlled in the acceptability rating experiment, the judges probably projected their own prosody onto the sentences and rendered a less common word order in a certain contrast condition acceptable.

Dutch speakers mostly produced an unaccented _nu_ when the object was contrasted but an L*H accented _nu_ when only the location was contrasted, which again poses the question whether this is simply a strategy to avoid an accent clash in two neighboring constituents. The co-occurrence patterns of the accentual realization of _nu_ and the object noun (Table 6) shows, however, that the combination of an accented _nu_ and an accented object is 3 times more likely than the combination between an unaccented _nu_ and an accented object. A further argument against the accent-clash-avoidance hypothesis is the observation that the intonational realization of _nu_ was independent of word order.

However, the results are not compatible with our earlier interpretation of the results from the English data that speakers mark whether the upcoming contrast is in the immediate constituent or further away (section 3.4). If this were the case, we would have expected an unaccented _nu_ for a contrast in location in the non-dominant word order. This is however not what we found. The Dutch data actually suggest that L*H is a device to prepare the listener for a less expected contrast, i.e., a contrast in a syntactically less salient, more oblique, constituent (the location), while an unaccented _nu_ signals a contrast in a syntactically more salient, less oblique, constituent (the object). This speculative connection between the obliqueness of syntactic constituents (Keenan & Comrie, 1977) and the degree to which participants expect a contrast in that constituent is supported by results from eye tracking experiments with these kinds of instructions, which have shown an initial bias towards a contrasting object (Dahan et al., 2002; Braun & Chen, 2008). Nevertheless, further empirical backing is needed to substantialize our speculation.

5 General Discussion

Despite the attested differences in word order flexibility of _NOW_ to signal upcoming contrasts, Dutch and English speakers’ prosodic realization of _NOW_ was remarkably similar. Neither group of speakers used phrasing to disambiguate the contrast conditions
(contrast in object, location or both) but both groups of speakers varied the intonational realization of now to mark the locus of the contrast. When the object was contrasted, the adverb was preferably unaccented or accented with a high level tone (H*) but when only the location was contrasted, the adverb was mostly realized with a low rise (L*H), in accordance with the patterns we observed in the experimental materials used by Dahan et al. (2002) and Chen et al. (2007).

The observed intonational patterns in NOW in the different contrast conditions cannot satisfactorily be explained by an accent-clash-avoidance strategy (not accenting NOW when the object is contrasted and therefore accented, accenting NOW when the object is given and therefore not accented). In fact, in English, an accented object was equally often preceded by an accented and by an unaccented NOW. In Dutch an accented object was three times as often preceded by an accented than by an unaccented nu. These data hence call for a functional, rather than a purely phonological explanation. One tentative explanation is elaborated in the following paragraph, drawing on observations on the interplay between word order and intonation.

Dutch speakers have the possibility to vary their word order for scope disambiguation, but in our study only a few speakers made use of word order variation. Strikingly, however, Dutch speakers’ choice of accent type on nu to signal the locus of the contrast — unaccented nu for a contrast in the object and an L*H accented nu for a contrast in the location — was independent of word order choice. This is an indication that the intonational realization of nu does not signal the immediacy of the contrast (whether it lies in the immediate constituent or in a constituent further away). Rather, it is conceivable that speakers produce nu with a comparatively prominent rising accent (L*H) to signal that the contrast lies in a less salient, more oblique, constituent (here, the location, cf. Keenan & Comrie 1977). On the other hand, speakers do not strongly mark nu with an accent if the contrast is found in a highly salient constituent (here, the object). This hypothesis ties in nicely with results from eye-tracking experiments showing that listeners tend to have a bias towards a contrast in the object (Dahan et al, 2002; Braun & Chen, 2008). However, as the experiments were not designed to test and compare these two interpretations and the sample of realizations with nu placed before the location was
small in the current study, further research is needed to give a definitive answer to the question whether the intonation of *nu* is sensitive to the obliqueness of the arguments.

Regarding the number of contrasts, English speakers and the majority of Dutch speakers did not vary the prosodic realization of *NOW* (phrasing and intonation) to disambiguate a single contrast in the object from a double contrast. This suggests that it is more important for speakers to mark whether the object is contrasted or not. Marking only whether there is a contrast in the object and not whether there is yet another contrast to come can be beneficial for both the listener and the speaker. For the listener, it would undoubtedly be difficult to distinguish three (or even more) meaningfully different pitch accents on such a short adverb. Furthermore, knowing whether to expect a contrast in the following constituent or not may speed up speech comprehension. For the speaker, it means that he or she need not plan the entire clause but may focus on the local structure. Whether there is an additional contrast later on may then be signalled by the prosodic realization of the object. This was indeed the case in our data. Dutch speakers, for instance, mostly produced a rising accent on the contrasted object when there was a further contrast in the location, but a fall when only the object was contrasted.

To conclude, the present data show that at least some Dutch speakers used word order to disambiguate the contrast conditions. Irrespective of their choice of word order, both Dutch speakers and English speakers varied the intonation of *NOW* to mark the upcoming contrasts. The intonation of *NOW* was primarily varied to signal the locus of the contrast (object vs. location), not the number of contrasts. This result was explained from a processing perspective. Surprisingly, differences in underlying word order flexibility did not influence the extent to which intonational variation in *NOW* was used. The exact choices of intonation patterns appear to be motivated by the salience of the contrasted constituent: *NOW* is unaccented if the contrast lies in a salient constituent (here the object), while it is accented with a prominent rise to signal a contrast in a less salient constituent (the location). This study extends our knowledge on the functions of the prosodic variations of the sequential adverb *now/nu* and adds to the body of literature on the interface between prosody and information structure, which has been mostly based on the prosodic realization of discourse referents (e.g., Brazil, 1975, Gussenhoven, 1984,
Pierrehumbert and Hirschberg, 1990; Steedman, 2000) and focus particles (e.g., Reis & Rosengren, 1997; Krifka, 1991).
Acknowledgements
We are grateful to Sam Hellmuth (University of York), Fang Liu (University of Chicago & University of London) for help with data collection. Further, we’d like to thank Alice Kruisselbrink and Marieke Hoetjes for making the movie clips, Mirjam Sparreboom for annotating the English and Dutch data at the word level, and Neil Forsyth for proofreading an earlier version of the manuscript.
Appendix

Please insert Table A1 here

Please insert Table A2 here

Please insert Table A3 here

Please insert Table A4 here

Please insert Table A5 here
References


Figure Captions

Figure 1. Box-plot showing the ratings of the English participants.

Figure 2. Box-plot showing the ratings of the Dutch participants.

Figure 3. Example displays of a DOUBLECONTRAST video clip in grayscale. Left panel shows the initial configuration, middle panel the display after the first trajectory, right panel the configuration after the second trajectory.

Figure 4. Pitch track of an utterance in ObjectContrast produced by an English speaker.

Figure 5. Pitch track of an utterance in LocationContrast produced by an English speaker.

Figure 6. Boxplot of the f0 excursion of the rise in semitones for the English accents (‘<’ indicates a significant difference at p < 0.025 between two accent types, while ‘=’ stands for ‘no difference’).

Figure 7. Pitch track of an utterance in ObjectContrast produced by a Dutch speaker.

Figure 8. Pitch track of an utterance in LocationContrast produced by a Dutch speaker.

Figure 9. Boxplot of the f0 excursion of the rise in semitones for Dutch pitch accent types (‘<’ indicates a significant difference at p < 0.016 between two accent types, while ‘=’ stands for ‘no difference’).

Figure 10. Boxplot of the f0 excursion of the fall in semitones for those Dutch pitch accents that did not differ in the f0 excursion of the rise (‘<’ indicates a significant difference at p < 0.025 between two accent types, while ‘=’ stands for ‘no difference’).
Tables

Table 1. Absolute and relative numbers of intonation patterns in the object noun in each condition in English

<table>
<thead>
<tr>
<th>Condition</th>
<th>!H*</th>
<th>!H*L</th>
<th>H*</th>
<th>H*L</th>
<th>L*</th>
<th>L*H</th>
<th>No accent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>LocationContrast</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>19</td>
<td>32</td>
</tr>
<tr>
<td>%</td>
<td>3.1%</td>
<td>6.2%</td>
<td>3.1%</td>
<td>3.1%</td>
<td>21.9%</td>
<td>3.1%</td>
<td>59.4%</td>
<td>100.0%</td>
</tr>
<tr>
<td>ObjectContrast</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>89.7%</td>
<td>3.4%</td>
<td>3.4%</td>
<td>3.4%</td>
<td>100.0%</td>
</tr>
<tr>
<td>DoubleContrast</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>23</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>%</td>
<td>0%</td>
<td>3.6%</td>
<td>10.7%</td>
<td>82.1%</td>
<td>0%</td>
<td>3.6%</td>
<td>0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>50</td>
<td>8</td>
<td>3</td>
<td>20</td>
<td>89</td>
</tr>
<tr>
<td>%</td>
<td>1.1%</td>
<td>3.4%</td>
<td>4.5%</td>
<td>56.2%</td>
<td>9.0%</td>
<td>3.4%</td>
<td>22.5%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
Table 2. The total number of occurrences and the mean % distribution of each intonation pattern of *now* in each condition in English

<table>
<thead>
<tr>
<th>Condition</th>
<th>H*</th>
<th>H*L</th>
<th>L*H</th>
<th>no accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>LocationContrast</td>
<td>9</td>
<td>4</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>%</td>
<td>28.1%</td>
<td>12.5%</td>
<td>31.3%</td>
<td>28.1%</td>
</tr>
<tr>
<td>ObjectContrast</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>%</td>
<td>28.1%</td>
<td>12.5%</td>
<td>0%</td>
<td>59.4%</td>
</tr>
<tr>
<td>DoubleContrast</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>%</td>
<td>15.6%</td>
<td>12.5%</td>
<td>6.3%</td>
<td>65.6%</td>
</tr>
</tbody>
</table>
Table 3. Co-occurrence of pitch accents of *now* (rows) and pitch accents in the object noun (columns)

<table>
<thead>
<tr>
<th>Pitch accent type of the object noun</th>
<th>!H*</th>
<th>!H*L</th>
<th>H*</th>
<th>H*L</th>
<th>L*</th>
<th>L*H</th>
<th>No accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>H*</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>H*L</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>L*H</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>No accent</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>28</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>50</td>
<td>8</td>
<td>3</td>
<td>20</td>
</tr>
</tbody>
</table>
Table 4. Absolute and relative numbers of intonation patterns in the object noun in each condition in Dutch

<table>
<thead>
<tr>
<th>Condition</th>
<th>!H*</th>
<th>!H*L</th>
<th>H*</th>
<th>H*L</th>
<th>L*H</th>
<th>No accent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>LocationContrast</td>
<td>N</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>14.3%</td>
<td>17.9%</td>
<td>21.4%</td>
<td>3.6%</td>
<td>25.0%</td>
<td>17.9%</td>
</tr>
<tr>
<td>ObjectContrast</td>
<td>N</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>17</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>0%</td>
<td>12.5%</td>
<td>8.3%</td>
<td>70.8%</td>
<td>8.3%</td>
<td>0%</td>
</tr>
<tr>
<td>DoubleContrast</td>
<td>N</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>7.1%</td>
<td>7.1%</td>
<td>17.9%</td>
<td>25.0%</td>
<td>35.7%</td>
<td>7.1%</td>
</tr>
<tr>
<td>Total</td>
<td>N</td>
<td>6</td>
<td>10</td>
<td>13</td>
<td>25</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>7.5%</td>
<td>12.5%</td>
<td>16.2%</td>
<td>31.2%</td>
<td>23.8%</td>
<td>8.8%</td>
</tr>
</tbody>
</table>
Table 5. The total number of occurrences and the mean % distribution of each intonation pattern of *nu* in each condition in Dutch

<table>
<thead>
<tr>
<th>Condition</th>
<th>H*</th>
<th>H*L</th>
<th>L*H</th>
<th>LH*</th>
<th>LH*L</th>
<th>No accent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location Contrast</strong></td>
<td>N</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>11.5%</td>
<td>3.1%</td>
<td>31.3%</td>
<td>40.5%</td>
<td>15.6%</td>
</tr>
<tr>
<td><strong>Object Contrast</strong></td>
<td>N</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>17.7%</td>
<td>3.1%</td>
<td>4.2%</td>
<td>14.3%</td>
<td>16.7%</td>
</tr>
<tr>
<td><strong>Double Contrast</strong></td>
<td>N</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>14.6%</td>
<td>14.6%</td>
<td>9.4%</td>
<td>23.8%</td>
<td>16.7%</td>
</tr>
</tbody>
</table>
Table 6: Co-occurrence of pitch accents of *nu* (rows) and pitch accents in the object noun (columns)

<table>
<thead>
<tr>
<th>Pitch accent type of the adverb</th>
<th>!H*</th>
<th>!H*L</th>
<th>H*</th>
<th>H*L</th>
<th>L*H</th>
<th>No accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>H*</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>H*L</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>L*H</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>LH*</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>LH*L</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>No accent</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>10</td>
<td>13</td>
<td>25</td>
<td>19</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 7. Intonation and phrasing of *nu* in non-dominant word order (Verb Noun nu Location)

<table>
<thead>
<tr>
<th>Rising accents</th>
<th>H*L</th>
<th>L*H</th>
<th>H*L H%</th>
<th>H*</th>
<th>No accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>LocationContrast</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ObjectContrast</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>DoubleContrast</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A1. Accent types used for the intonational annotation of the adverb. The square in the stylization indicates the syllable. Standard IViE accent types, which were used for the annotation of British English *now*, are marked in bold face.

<table>
<thead>
<tr>
<th>Accent type</th>
<th>Stylized contour</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H*</td>
<td></td>
<td>High tone in <em>now</em>, preceded by a weak rise, followed by a high level stretch</td>
</tr>
<tr>
<td>H*L</td>
<td></td>
<td>High tone in <em>now</em>, preceded by a weak rise, followed by a sharp pitch fall within the stressed syllable to the next one</td>
</tr>
<tr>
<td>LH*</td>
<td></td>
<td>Same as H* but preceded by a noticeable f0-rise</td>
</tr>
<tr>
<td>LH*L</td>
<td></td>
<td>Same as H*L but preceded by a noticeable f0-rise</td>
</tr>
<tr>
<td>L*H</td>
<td></td>
<td>Low tone in <em>now</em> that is low in the speaker’s register, followed by a steep f0-rise</td>
</tr>
</tbody>
</table>
Table A2. Mean (and sd) of acoustic measurements the pitch accent types of *now*

<table>
<thead>
<tr>
<th></th>
<th>H*</th>
<th>H*L</th>
<th>L*H</th>
<th>No accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0 excursion of the rise in st</td>
<td>0.23</td>
<td>0.99</td>
<td>2.64</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(1.19)</td>
<td>(1.49)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>F0 excursion of the fall in st</td>
<td>0.95</td>
<td>4.81</td>
<td>0.50</td>
<td>1.90</td>
</tr>
<tr>
<td></td>
<td>(0.81)</td>
<td>(1.81)</td>
<td>(0.60)</td>
<td>(1.55)</td>
</tr>
<tr>
<td>Relative duration of nu (divided by verb duration)</td>
<td>1.30</td>
<td>1.81</td>
<td>1.27</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.65)</td>
<td>(0.52)</td>
<td>(0.35)</td>
</tr>
<tr>
<td>Alignment of the f0-maximum with respect to syllable onset in ms</td>
<td>120.6</td>
<td>87.0</td>
<td>206.3</td>
<td>45.2</td>
</tr>
<tr>
<td></td>
<td>(145.4)</td>
<td>(49.1)</td>
<td>(63.4)</td>
<td>(73.5)</td>
</tr>
<tr>
<td>Alignment of the f0-minimum with respect to syllable onset in ms</td>
<td>109.5</td>
<td>262.1</td>
<td>40.4</td>
<td>158.2</td>
</tr>
<tr>
<td></td>
<td>(70.7)</td>
<td>(47.6)</td>
<td>(21.4)</td>
<td>(75.6)</td>
</tr>
</tbody>
</table>
Table A3. Mean (and sd) of phonetic variables for the different Dutch pitch accent types

<table>
<thead>
<tr>
<th></th>
<th>H*</th>
<th>H*L</th>
<th>L*H</th>
<th>LH*</th>
<th>LH*L</th>
<th>No accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0 excursion of the rise in st</td>
<td>1.86</td>
<td>0.99</td>
<td>5.84</td>
<td>2.91</td>
<td>3.17</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>(0.51)</td>
<td>(16)</td>
<td>(2.16)</td>
<td>(0.92)</td>
<td>(1.03)</td>
<td>(0.74)</td>
</tr>
<tr>
<td>F0 excursion of the fall in st</td>
<td>1.05</td>
<td>2.54</td>
<td>3.58</td>
<td>0.78</td>
<td>2.34</td>
<td>1.90</td>
</tr>
<tr>
<td></td>
<td>(0.91)</td>
<td>(0.95)</td>
<td>(3.96)</td>
<td>(1.59)</td>
<td>(1.03)</td>
<td>(1.55)</td>
</tr>
<tr>
<td>Relative duration of nu (divided by verb duration)</td>
<td>0.31</td>
<td>0.27</td>
<td>0.38</td>
<td>0.32</td>
<td>0.29</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.06)</td>
<td>(0.14)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Alignment of the f0-maximum with respect to syllable onset in ms</td>
<td>52.7</td>
<td>56.7</td>
<td>181.4</td>
<td>141.3</td>
<td>97.1</td>
<td>33.2</td>
</tr>
<tr>
<td></td>
<td>(82.7)</td>
<td>(57.2)</td>
<td>(66.8)</td>
<td>(33.6)</td>
<td>(44.1)</td>
<td>(50.3)</td>
</tr>
<tr>
<td>Alignment of the f0-minimum with respect to syllable onset in ms</td>
<td>56.3</td>
<td>94.2</td>
<td>35.9</td>
<td>32.4</td>
<td>46.4</td>
<td>111.8</td>
</tr>
<tr>
<td></td>
<td>(41.5)</td>
<td>(76.6)</td>
<td>(37.5)</td>
<td>(20.4)</td>
<td>(60.5)</td>
<td>(44.3)</td>
</tr>
</tbody>
</table>
Table A4. Intonation of *nu* in the dominant word order from speakers in the *variable* word order group

<table>
<thead>
<tr>
<th></th>
<th>H*</th>
<th>H*L</th>
<th>L*H</th>
<th>LH*</th>
<th>No accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>LocationContrast</td>
<td>4</td>
<td>1</td>
<td>17</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>ObjectContrast</td>
<td>9</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>DoubleContrast</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>
Table A5. Position of *nu* in intonation phrase (IP) in non-dominant word order (e.g. *Verplaats het boek nu naar vak 3*)

<table>
<thead>
<tr>
<th>LocationContrast</th>
<th>break after nu</th>
<th>break before nu</th>
<th>No break</th>
</tr>
</thead>
<tbody>
<tr>
<td>ObjectContrast</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>DoubleContrast</td>
<td>5</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>
Figure

```
| now | put | the | bell | in | cell | five |
| %L | H*L | H*L | L%   |
```

```
Time (s)
```

0.6797 - 2.982
now put the banana into cell one
Figure

Pitch accent type

F0 excursion of the rise (st)
<table>
<thead>
<tr>
<th>Time (s)</th>
<th>0</th>
<th>2.297</th>
</tr>
</thead>
</table>

---

**Figure**

<table>
<thead>
<tr>
<th>verplaats</th>
<th>nu</th>
<th>de</th>
<th>gitaar</th>
<th>naar</th>
<th>vak</th>
<th>een</th>
</tr>
</thead>
<tbody>
<tr>
<td>move</td>
<td>now</td>
<td>the</td>
<td>guitar</td>
<td>into</td>
<td>cell</td>
<td>one</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>%L</th>
<th>H*</th>
<th>H*L</th>
<th>L%</th>
</tr>
</thead>
</table>


now put the bell in cell five

now put the banana into cell one

verplaats nu de deur naar vak negen
move now the door into cell nine

Time (s)
Figure

Pitch accent type

F0 excursion of the rise (st)
Pitch accent type

F0 excursion of the fall (st)