Frequency Matters: Pitch Accents and Information Status

Katrin Schweitzer, Michael Walsh, Bernd Möbius,
Arndt Riester, Antje Schweitzer, Hinrich Schütze
University of Stuttgart
Stuttgart, Germany
<firstname>.<surname>@ims.uni-stuttgart.de

Abstract

This paper presents the results of a series of experiments which examine the impact of two information status categories (*given* and *new*) and frequency of occurrence on pitch accent realisations. More specifically the experiments explore within-type similarity of pitch accent productions and the effect information status and frequency of occurrence have on these productions. The results indicate a significant influence of both pitch accent type and information status category on the degree of within-type variability, in line with exemplar-theoretic expectations.

1 Introduction

It seems both intuitive and likely that prosody should have a significant role to play in marking information status in speech. While there are well established expectations concerning typical associations between categories of information status and categories of pitch accent, e.g. rising L∗H accents are often a marker for givenness, there is nevertheless some variability here (Baumann, 2006). Furthermore, little research has focused on how pitch accent tokens of the same type are realised nor have the effects of information status and frequency of occurrence been considered.

From the perspective of speech technology, the tasks of automatically inferring and assigning information status clearly have significant importance for speech synthesis and speech understanding systems.

The research presented in this paper examines a number of questions concerning the relationship between two information status categories (*new* and *given*), and how tokens of associated pitch accent types are realised. Furthermore the effect of frequency of occurrence is also examined from an exemplar-theoretic perspective.

The questions directly addressed in this paper are as follows:

1. How are different tokens of a pitch accent type realised?
   Does frequency of occurrence of the pitch accent type play a role?
2. What effect does information status have on realisations of a pitch accent type?
   Does frequency of occurrence of the information status category play a role?
3. Does frequency of occurrence in pitch accents *and* in information status play a role, i.e. is there a combined effect?

In examining the realisation of pitch accent tokens, their degree of similarity is the characteristic under investigation. Similarity is calculated by determining the cosine of the angle between pairs of pitch accent vector representations (see section 6).

The results in this study are examined from an exemplar-theoretic perspective (see section 3). The expectations within that framework are based upon two different aspects. Firstly, it is expected that, since all exemplars are stored, exemplars of a type that occur often, offer the speaker a wider selection of exemplars to choose from during production (Schweitzer and Möbius, 2004), i.e. the realisations are expected to be more variable than those of a rare type. However, another aspect of Exemplar Theory has to be considered, namely entrenchment (Pierrehumbert, 2001; Bybee, 2006). The central idea here is that frequently occurring behaviours undergo processes of entrenchment, they become in some sense routine. Therefore realisations of a very frequent type are expected to be realised similar to each other. Thus, similarity and variability are expressions of the same characteristic: the higher the degree of similarity of pitch accent tokens, the lower their realisation variability.

Proceedings of the 12th Conference of the European Chapter of the ACL, pages 728–736,
Athens, Greece, 30 March – 3 April 2009.

The structure of this paper is as follows: Section 2 briefly examines previous work on the interaction of information status categories and pitch accents. Section 3 provides a short introduction to Exemplar Theory. In this study similarity of pitch accent realisations on syllables, annotated with the information status categories of the words they belong to, is examined using the parametric intonation model (Möhler, 1998) which is outlined in Section 4. Section 5 discusses the corpus employed. Section 6 introduces a general methodology which is used in the experiments in Sections 7, 8 and 9. Section 10 then presents some discussion, conclusions and opportunities for future research.

2 Information Status and Intonation

It is commonly assumed that pitch accents are the main correlate of information status[1] in speech (Halliday, 1967). Generally, accenting is said to signal novelty while deaccenting signals given information (Brown, 1983), although there is counter evidence: various studies note given information being accented (Yule, 1980; Bard and Aylett, 1999). Terken and Hirschberg (1994) point out that new information can also be deaccented.

As for the question of which pitch accent type (in terms of ToBI categories (Silverman et al., 1992)) is typically assigned to different degrees of givenness, Pierrehumbert and Hirschberg (1990) find H$*$ to be the standard novelty accent for English, a finding which has also been confirmed by Baumann (2006) and Schweitzer et al. (2008) for German. Given information on the other hand, if accented at all, is found to carry L$*$ accent in English (Pierrehumbert and Hirschberg, 1990). Baumann (2006) finds deaccentuation to be the most preferred realisation for givenness in his experimental phonetics studies on German. However, Baumann (2006) points out that H+L$*$ has also been found as a marker of givenness in a German corpus study. Previous findings on the corpus used in the present study found L$*$H being the typical marker for givenness (Schweitzer et al., 2008).

Leaving the phonological level and examining correlates of information status in acoustic detail, Kohler (1991) reports that in a falling accent, an early peak indicates established facts, while a medial peak is used to mark novelty. In a recent study Kügler and Féry (2008) found givenness to lower the high tones of prenuclear pitch accents and to cancel them out postnuclearly. These findings among others (Kügler and Féry, 2008) motivate an examination of the acoustic detail of pitch accent shape across different information status categories.

The experiments presented here go one step further, however, in that they also investigate potential exemplar-theoretic effects.

3 Exemplar Theory

Exemplar Theory is concerned with the idea that the acquisition of language is significantly facilitated by repeated exposure to concrete language input, and it has successfully accounted for a number of language phenomena, including diachronic language change and frequency of occurrence effects (Bybee, 2006), the emergence of grammatical knowledge (Abbot-Smith and Tomasello, 2006), syllable duration variability (Schweitzer and Möbius, 2004; Walsh et al., 2007), entrenchment and lenition (Pierrehumbert, 2001), among others. Central to Exemplar Theory are the notions of exemplar storage, frequency of occurrence, recency of occurrence, and similarity. There is an increasing body of evidence which indicates that significant storage of language input exemplars, rich in detail, takes place in memory (Johnson, 1997; Croot and Rastle, 2004; Whiteside and Varley, 1998). These stored exemplars are then employed in the categorisation of new input percepts. Similarly, production is facilitated by accessing these stored exemplars. Computational models of the exemplar memory also argue that it is in a constant state of flux with new inputs updating it and old unused exemplars gradually fading away (Pierrehumbert, 2001).

Up to now, virtually no exemplar-theoretic research has examined pitch accent prosody (but see Marsi et al. (2003) for memory-based prediction of pitch accents and prosodic boundaries, and Walsh et al. (2008)(discussed below)) and to the authors' knowledge this paper represents the first attempt to examine the relationship between pitch accent prosody and information status from an exemplar-theoretic perspective. Given the considerable weight of evidence for the influence of frequency of occurrence effects in a variety of other linguistic domains it seems reasonable to explore such effects on pitch accent and information sta-

[1] The term *information status* is used in (Prince, 1992) for the first time. Before that the terms *givenness*, *novelty* or *information structure* were used for these concepts.

tus realisations. For example, what effect might givenness have on a frequently/infrequently occurring pitch accent? Does novelty produce a similar result?

The search for possible frequency of occurrence effects takes place with respect to pitch accent shapes captured by the parametric intonation model discussed next.

4 The Parametric Representation of Intonation Events - PaIntE

The model approximates stretches of F_0 by employing a phonetically motivated model function (Möhler, 1998). This function consists of the sum of two sigmoids (rising and falling) with a fixed time delay which is selected so that the peak does not fall below 96% of the function's range. The resulting function has six parameters which describe the contour and were employed in the analysis: parameters a1 and a2 express the gradient of the accent's rise and fall, parameter b describes the accent's temporal alignment (which has been shown to be crucial in the description of an accent's shape (van Santen and Möbius, 2000)), c1 and c2 model the ranges of the rising and falling amplitude of the accent's contour, respectively, and parameter d expresses the peak height of the accent.[2] These six parameters are thus appropriate to describe different pitch accent shapes.

For the annotation of intonation the GToBI(S) annotation scheme (Mayer, 1995) was used. In earlier versions of PaIntE, the approximation of the F_0-contour for H∗L and H∗ was carried out on the accented and post–accented syllables. However, for these accents the beginning of the rise is likely to start at the preaccented syllable. In the current version of PaIntE the window used for the approximation of the F_0-contour for H∗L and H∗ accents has been extended to the preaccented syllable, so that the parameters are calculated over the span of the accented syllables and its immediate neighbours (unless it is followed by a boundary tone which causes the window to end at the end of the accented syllable).

5 Corpus

The experiments that follow (sections 7, 9 and 8), were carried out on German pitch accents from the IMS Radio News Corpus (Rapp, 1998). This corpus was automatically segmented and manually labelled according to GToBI(S) (Mayer, 1995). In the corpus, 1233 syllables are associated with an L∗H accent, 704 with an H∗L accent and 162 with an H∗ accent.

The corpus contains data from three speakers, two female and a male one, but the majority of the data is produced by the male speaker (888 L∗H accents, 527 H∗L accents and 152 H∗ accents). In order to maximise the number of tokens, all three speakers were combined. Of the analysed data, 77.92% come from the male speaker. However, it is not necessarily the case that the same percentage of the variability also comes from this speaker: Both, PaIntE and z-scoring (cf. section 6) normalise across speakers, so the contribution from each individual speaker is unclear.

The textual transcription of the corpus was annotated with respect to information status using the annotation scheme proposed by Riester (2008). In this taxonomy information status categories reflect the default contexts in which presuppositions are resolved, which include e. g. discourse context, environment context or encyclopaedic context.

The annotations are based solely on the written text and follow strict semantic criteria. Given that textual information alone (i.e. without prosodic or speech related information) is not necessarily sufficient to unambiguously determine the information status associated with a particular word, there are therefore cases where words have multiple annotations, reflecting underspecification of information status. However, it is important to note that in all the experiments reported here, only unambiguous cases are considered.

The rich annotation scheme employed in the corpus makes establishing inter-annotator agreement a time-consuming task which is currently underway. Nevertheless, the annotation process was set up in a way to ensure a maximal smoothing of uncertainties. Texts were independently labelled by two annotators. Subsequently, a third, more experienced annotator compared the two results and, in the case of discrepancies, took a final decision.

In the present study the categories *given* and *new* are examined. These categories do not represent a binary distinction but are two extremes from a set of clearly distinguished categories. For the most part they correspond to the categories *textually given* and *brand-new* that are used in Bau-

[2] Further information and illustrations concerning the mechanics of the PaIntE model can be found in Möhler and Conkie (1998).

mann (2006), but their scope is more tightly constrained. The information status annotations are mapped to the phonetically transcribed speech signals, from which individual syllable tokens bearing information status are derived.

Syllables for which one of the PaIntE-parameters was identified as an outlier, were removed. Outliers were defined such that the upper 2.5 percentile as well as the lower 2.5 percentile of the data were excluded. This led to a reduced number of pitch accent tokens: 1021 L∗H accents, 571 H∗L accents and 134 H∗ accents. Thus, there is a continuum of frequency of occurrence, high to low, from L∗H to H∗.

With respect to information status, 102 L∗H accents, 87 H∗L accents and 21 H∗ accents were unambiguously labelled as *new*. For givenness the number of tokens is: 114 L∗H accents, 44 H∗L accents and 10 H∗ accents.

6 General Methodology

In the experiments the general methodology for calculation of similarity detailed in this section was employed.

For tokens of the pitch accent types L∗H, H∗L and H∗, each token was modelled using the full set of PaIntE parameters. Thus, each token was represented in terms of a 6-dimensional vector.

For each of the pitch accent types the following steps were carried out:

– For each 6-dimensional pitch accent category token calculate the z-score value for each dimension. The z-score value represents the number of standard deviations the value is away from the mean value for that dimension and allows comparison of values from different normal distributions. The z-score is given by:

$$z - score_{dim} = \frac{value_{dim} - mean_{dim}}{sdev_{dim}} \quad (1)$$

Hence, at this point each pitch accent is represented by a 6-dimensional vector where each dimension value is a z-score.

– For each token z-scored vector calculate how similar it is to every other z-scored vector within the same pitch accent category, and, in Experiment 2 and 3, with the same information status value (e.g. new), using the cosine of the angle between the vectors. This is given by:

$$cos(\vec{i}, \vec{j}) = \frac{\vec{i} \bullet \vec{j}}{\| \vec{i} \| \| \vec{j} \|} \quad (2)$$

where *i* and *j* are vectors of the same pitch accent category and • represents the dot product.

Each comparison between vectors yields a similarity score in the range [-1,1], where -1 represents high dissimilarity and 1 represents high similarity.

The experiments that follow examine distributions of token similarity. In order to establish whether distributions differ significantly two different levels of significance were employed, depending on the number of pairwise comparisons performed.

When comparing two distributions (i.e. performing one test), the significance level was set to $\alpha = 0.05$. In those cases where multiple tests were carried out (Experiment 1 and Experiment 3), the level of significance was adjusted (Bonferroni correction) according to the following formula:

$$\alpha = 1 - (1 - \alpha_1)^{\frac{1}{n}} \quad (3)$$

where α_1 represents the target significance level (set to 0.05) and n represents the number of tests being performed. The Bonferroni correction is often discussed controversially. The main criticism concerns the increased likelihood of type II errors that lead to non-significance of actually significant findings (Pernegger, 1998). Although this conservative adjustment was applied, the statistical tests in this study resulted in significant p-values indicating the robustness of the findings.

7 Experiment 1: Examining frequency of occurrence effects in pitch accents

In accordance with the general methodology set out in section 6, the PaIntE vectors of pitch accent tokens of types L∗H, H∗L, and H∗ were all z-scored and, within each type, every token was compared for similarity against every other token of the same type, using the cosine of the angle between their vectors. In essence, this experiment illustrates how similarly pitch accents of the same type are realised.

Figure 1 depicts the results of the analysis. It shows the density plot for each distribution of cosine-similarity comparison values, whereby the

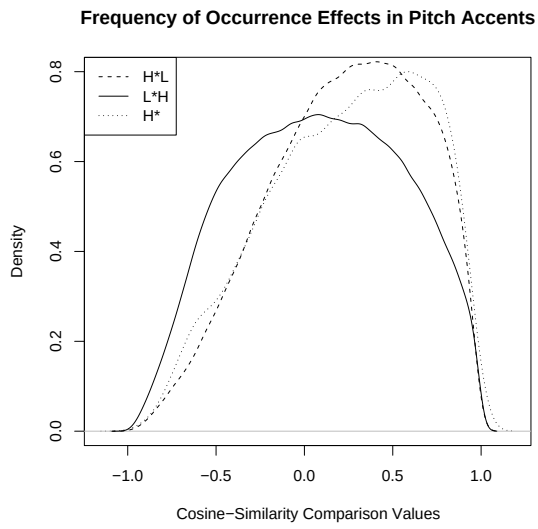

Figure 1: *Density plots for similarity within pitch accent types. All distributions differ significantly from each other. There is a trend towards greater similarity from high-frequency* L∗H *to low-frequency* H∗.

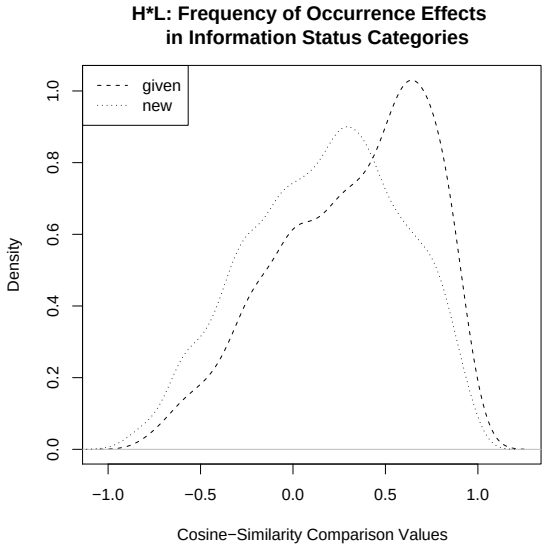

Figure 2: *Density plots for similarity of* H∗L *tokens. Tokens of the low-frequency information status category* given *display greater similarity to each other than those of the high-frequency information status category* new.

distributions can be compared directly – irrespective of the different number of data points.

An initial observation is that L∗H tokens tend to be realised fairly variably, the main portion of the distribution is centred around zero. Tokens of H∗L tend to be produced more similarly (i.e. the distribution is centred around a higher similarity value), and tokens of H∗ more similarly again. These three distributions were tested against each other for significance using the Kolmogorov-Smirnov test ($\alpha = 0.017$), yielding p-values of $p \ll 0.001$. Thus there are significant differences between these distributions.

What is particularly noteworthy is that a *decrease* in frequency of occurrence across pitch accent types co-occurs significantly with an *increase* in within-type token similarity.

While the differences between the graphed distributions do not appear to be highly marked the frequency of occurrence effect is nevertheless in keeping with exemplar-theoretic expectations as posited by Bybee (2006) and Schweitzer and Möbius (2004), that is, the high frequency of occurrence entails a large number of stored exemplars, giving the speaker the choice from among a large number of production targets. This wider choice leads to a broader range of chosen targets for different productions and thus to more variable realisations of tokens of the same type.

Walsh et al. (2008) also reported significant differences between these distributions, however, there did not appear to be a clear frequency of occurrence effect. The results in the present study differ from their results because the distributions centre around different ranges of the similarity scale clearly indicating that each accent type behaves differently in terms of similarity/variability between the tokens of the respective type. The differences between the two findings can be ascribed to the augmented PaIntE model (section 4).

Given the results from this experiment, the next experiment seeks to establish what relationship, if any, exists between information status and pitch accent production variability.

8 Experiment 2: Examining frequency of occurrence effects in information status categories

This experiment was carried out in the same manner as Experiment 1 above with the exception that in this experiment a subset of the corpus was employed: only syllables that were unambiguously labelled with either the information status category *new* or the category *given* were included in the analyses. The experiment aims to investigate the effect of information status on the similarity/variability of tokens of different pitch accent types. For each pitch accent type, tokens that were labelled with the information status category *new*

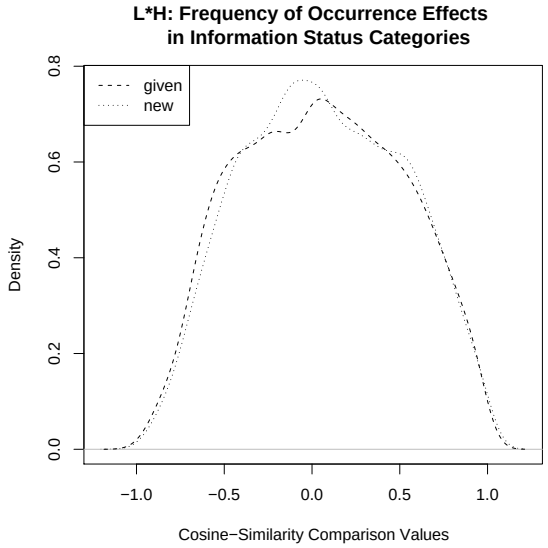

Figure 3: *Density plots for similarity of* L∗H *tokens. The curves differ significantly, a trend towards greater similarity is not observable. The number of tokens for both information status categories is comparable.*

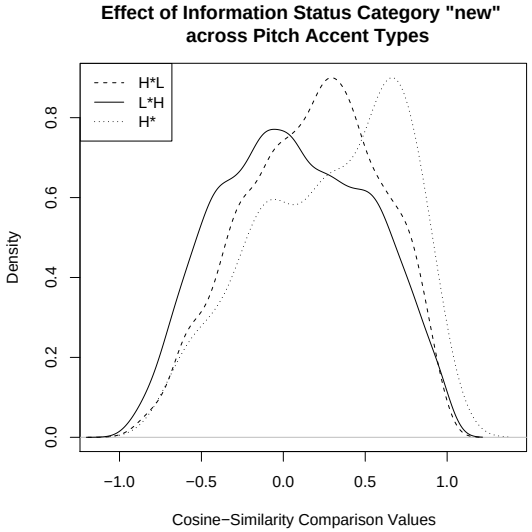

Figure 4: *Density plots for similarity of* new *tokens across three pitch accent types. In comparison to fig. 1 the trend towards greater similarity from high-frequency* L∗H *to low-frequency* H∗ *is even more pronounced.*

were compared to tokens labelled as *given*. Again, a pairwise Kolmogorov-Smirnov test was applied for each comparison ($\alpha = 0.05$). Figure 2 depicts the results for H∗L accents. The K-S test yielded a highly significant difference between the two distributions ($p \ll 0.001$), reflecting the clearly visible difference between the two curves. It is noteworthy here that for H∗L the information status category *new* is more frequent than the category *given*. Indeed, approximately twice as many are labelled as *new* than those labelled *given*. Figure 2 illustrates that *new* H∗L accents are realised more variably than *given* ones. That is, again, an increase in frequency of occurrence co-occurs with an increase in similarity, this time at the level of information status.

Figure 3 depicts the difference in similarity/variability for L∗H between *new* tokens and *given* tokens. It is clearly visible that the two curves do not differ as much as those under the H∗L condition. Both curves centre around zero reflecting the fact that for both types the tokens are variable. Although the Kolmogorov-Smirnov test indicates significance ($\alpha = 0.05, p = 0.044$), the nature of the impact that information status has in this case is unclear.

Here again an effect of frequency of occurrence might be the reason for this result. The high frequency of L∗H accents in general results in a relative high frequency of *given* L∗H tokens. So the token number for both types is similar (102 *new* L∗H tokens vs. 114 *given* L∗H tokens), there is high frequency in both cases, hence variability.

These results, particularly in the case of H∗L (fig. 2) indicate that information status affects pitch accent realisation. The next experiment compares the effect across different pitch accent types.

9 Experiment 3: Examining the effect of information status across pitch accent types

This experiment was carried out in the same manner as Experiments 1 and 2 above. For each pitch accent type, figure 4 depicts within-type pitch accent similarity for tokens unambiguously labelled as *new*.

As with Experiments 1 and 2, frequency of occurrence once more appears to play a significant role. Again, all Kolmogorov-Smirnov tests yielded significant results ($p < 0.017$ in all cases). Indeed, the difference between the distributions of L∗H, H∗L, and H∗ similarity plots appears to be considerably more prominent than in Experiment 1 (see fig. 1). This indicates that under the condition of *novelty* the frequency of occurrence effect is more pronounced. In other words, there is a considerably more noticeable difference across the distributions of L∗H, H∗L and H∗, when nov-

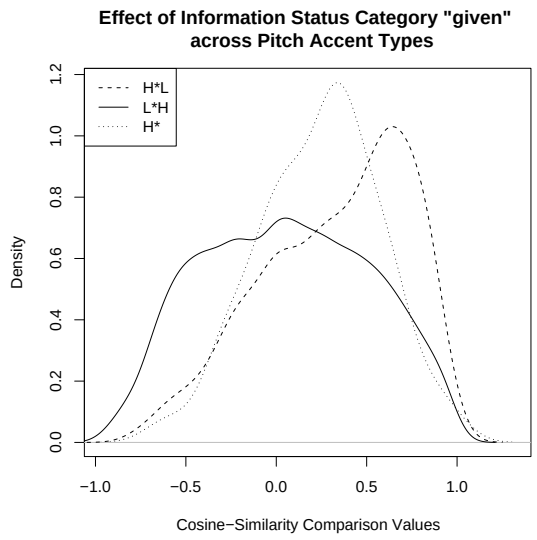

Figure 5: *Density plots for similarity of* given *tokens across three pitch accent types. Mid-frequency* H∗L *displays greater similarity than high-frequency* L∗H*. For lowest frequency* H∗ *(only 10 tokens) the trend cannot be observed.*

elty is considered: novelty compounds the frequency of occurrence effect.

Figure 5 illustrates results of the same analysis methodology but applied to tokens of pitch accents unambiguously labelled as *given*. Once again there is a considerable difference between the distributions of L∗H and H∗L tokens ($p < 0.017$). And again, this difference reflects a more pronounced frequency of occurrence effect for *given* tokens than for all accents pooled (as described in Experiment 1): the information status category *given* compounds the frequency of occurrence effect for L∗H and H∗L.

For H∗ the result is not as clear as for the two more frequent accents. The comparison between H∗ and L∗H results in a significant difference ($p < 0.017$) whereas the comparison between H∗ and H∗L is slightly above the conservative significance level ($p = 0.0186$). Moreover, the distribution is centred between the distributions for L∗H and H∗L and it is thus not clear how to interpret this result with respect to a possible frequency of occurrence effect. However, having only ten instances of *given* H∗, the explanatory power of these comparisons is questionable.

10 Discussion

The experiments discussed above yield a number of interesting results with implications for research in prosody, information status, the interaction between the two domains, and for exemplar theory.

Returning to the first question posed at the outset in section 1, it is quite clear from Experiment 1 that a certain amount of variability exists when different tokens of the same pitch accent type are produced. It is also clear, from the same experiment, that the frequency of occurrence of the pitch accent type does indeed play a role: with an increase in frequency comes an increase in variability. This result is in line with the exemplar-theoretic view that since all exemplars are stored, exemplars of a type that occur often are more variable because they offer the speaker a wider selection of exemplars to choose from during production (Schweitzer and Möbius, 2004). However, with respect to entrenchment (Pierrehumbert, 2001; Bybee, 2006), i.e. the idea that frequently occurring behaviours undergo processes of entrenchment, in Experiment 1 one might expect to see greater similarity in the realisations of L∗H. However, it is important to note that while tokens of L∗H are not particularly similar to each other (the bulk of the distribution is around zero (see figure 1)), they are not too dissimilar either. That is, they rest at the midpoint of the similarity continuum produced by cosine calculation, in quite a normal looking distribution. This is not at odds with the idea of entrenchment. As productions of a pitch accent type become more frequent, the distribution of similarity spreads from the right side of the graph (where infrequent and highly similar H∗ tokens lie) leftwards (through H∗L) to the point where the L∗H distribution is found. Beyond this point tokens are excessively different.

The second question posed in section 1, and addressed in Experiment 2, sought to ascertain the impact, if any, information status has on pitch accent realisation. Distributions of *given* and *new* H∗L similarity scores differed significantly, as did distributions of *given* and *new* L∗H similarity scores, indicating that information status affects realisation. In other words, for both pitch accent types, *given* and *new* tokens behave differently. Concerning the frequency of occurrence of the information status categories, certainly in the case of H∗L the higher frequency *new* tokens exhibited more variability. In the case of L∗H similar numbers of *new* and *given* tokens, possibly due to the high frequency of L∗H in general,

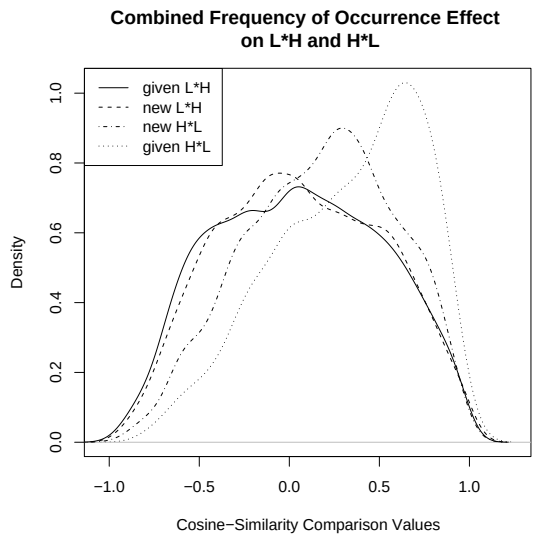

Figure 6: *Density plots for similarity of combinations of information status categories* given *and* new *with pitch accent types* L∗H *and* H∗L. *The distributions show a clear trend towards greater similarity form high-frequency "*given L∗H*" and "*new L∗H*" to mid-frequency "*new H∗L*" and low-frequency "*given H∗L*".*

led to visually similar yet significantly different distributions. Once again sensitivity to frequency of occurrence seems to be present, in line with exemplar-theoretic predictions.

The final question concerns the possibility of a combined effect of pitch accent frequency of occurrence and information status frequency of occurrence. Figures 4 and 5 depict a clear compounding effect of both information status categories across the different pitch accent types (and their inherent frequencies) when compared to figure 1. Interestingly, the less frequently occurring *given* appears to have a greater impact, particularly on high frequency L∗H.

Figure 6 displays all possible combinations of L∗H, H∗L, *given* and *new*. H∗ is omitted in this graph because of the small number of tokens (10 *given*, 21 *new*) and the resulting lack of explanatory power. It is evident that an overall frequency of occurrence effect can be observed: "*given* L∗H" and "*new* L∗H", which have a similar number of instances (114 vs. 102 tokens) both centre around zero and are thus the most leftward skewed curves in the graph. The distribution of "*new* H∗L" (87 tokens) shows a trend towards the right hand side of the graph and thus represents greater similarity of the tokens. The distribution of similarity values for the least frequent combination of pitch accent and information status, "*given* H∗L" (44 tokens), centres between 0.5 and 1.0 and is thus the most rightward curve in the graph, reflecting the highest similarity between the tokens.

These results highlight an intricate relationship between pitch accent production and information status. The information status of the word influences not only the type and shape of the pitch accent (Pierrehumbert and Hirschberg, 1990; Baumann, 2006; Kügler and Féry, 2008; Schweitzer et al., 2008) but also the similarity of tokens within a pitch accent type. Moreover, this effect is well explainable within the framework of Exemplar Theory as it is subject to frequency of occurrence: tokens of rare types are produced more similar to each other than tokens of frequent types.

In the context of speech technology, unfortunately the high variability in highly frequent pitch accents has a negative consequence, as the correlation between a certain pitch accent or a certain information status category and the F_0 contour is not a one-to-one relationship. However, forewarned is forearmed and perhaps a finer grained contextual analysis might yield more context specific solutions.

11 Future Work

The methodology outlined in section 6 gives a lucid insight into the levels of similarity found in pitch accent realisations. Further insights, however, could be gleaned from a fine-grained examination of the PaIntE parameters. For example, which parameters differ and under what conditions when examining highly variable tokens? Information status evidently plays a role in pitch accent production but the contexts in which this takes place have yet to be examined. In addition, the role of information *structure* (focus-background, contrast) also needs to be investigated. A further line of research worth pursuing concerns the impact of information status on the temporal structure of spoken utterances and possible compounding with frequency of occurrence effects.

References

Kirsten Abbot-Smith and Michael Tomasello. 2006. Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, 23(3):275–290.

Ellen G. Bard and M. P. Aylett. 1999. The dissociation of deaccenting, givenness, and syntactic role in

spontaneous speech. In *Proceedings of ICPhS (San Francisco)*, volume 3, pages 1753–1756.

Stefan Baumann. 2006. *The Intonation of Givenness – Evidence from German.*, volume 508 of *Linguistische Arbeiten*. Niemeyer, Tübingen. Ph.D. thesis, Saarland University.

Gillian Brown. 1983. Prosodic structure and the given/new distinction. In Anne Cutler and D. Robert Ladd, editors, *Prosody: Models and Measurements*, pages 67–77. Springer, New York.

Joan Bybee. 2006. From usage to grammar: The mind's response to repetition. *Language*, 84:529–551.

Karen Croot and Kathleen Rastle. 2004. Is there a syllabary containing stored articulatory plans for speech production in English? In *Proceedings of the 10th Australian International Conference on Speech Science and Technology (Sydney)*, pages 376–381.

Michael A. K. Halliday. 1967. *Intonation and Grammar in British English*. Mouton, The Hague.

Keith Johnson. 1997. Speech perception without speaker normalization: An exemplar model. In K. Johnson and J. W. Mullennix, editors, *Talker Variability in Speech Processing*, pages 145–165. Academic Press, San Diego.

Klaus J. Kohler. 1991. Studies in german intonation. *AIPUK (Univ. Kiel)*, 25.

Frank Kügler and Caroline Féry. 2008. Pitch accent scaling on given, new and focused constituents in german. *Journal of Phonetics*.

Erwin Marsi, Martin Reynaert, Antal van den Bosch, Walter Daelemans, and Véronique Hoste. 2003. Learning to predict pitch accents and prosodic boundaries in dutch. In *Proceedings of the ACL-2003 Conference (Sapporo, Japan)*, pages 489–496.

Jörg Mayer. 1995. Transcribing German Intonation – The Stuttgart System. Technical report, Universität Stuttgart. http://www.ims.uni-stuttgart.de/phonetik/joerg/labman/STGTsystem.html.

Gregor Möhler and Alistair Conkie. 1998. Parametric modeling of intonation using vector quantization. In *Third Intern. Workshop on Speech Synth (Jenolan Caves)*, pages 311–316.

Gregor Möhler. 1998. Describing intonation with a parametric model. In *Proceedings ICSLP*, volume 7, pages 2851–2854.

T. V. Pernegger. 1998. What's wrong with Bonferroni adjustment. *British Medical Journal*, 316:1236–1238.

Janet Pierrehumbert and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 271–311. MIT Press, Cambridge.

Janet Pierrehumbert. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee and Paul Hopper, editors, *Frequency and the Emergence of Linguistic Structure*, pages 137–157. Amsterdam.

Ellen F. Prince. 1992. The ZPG Letter: Subjects, Definiteness and Information Status. In W. C. Mann and S. A. Thompson, editors, *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text*, pages 295–325. Amsterdam.

Stefan Rapp. 1998. *Automatisierte Erstellung von Korpora für die Prosodieforschung*. Ph.D. thesis, IMS, Universität Stuttgart. AIMS 4 (1).

Arndt Riester. 2008. A Semantic Explication of Information Status and the Underspecification of the Recipients' Knowledge. In Atle Grønn, editor, *Proceedings of Sinn und Bedeutung 12*, Oslo.

Antje Schweitzer and Bernd Möbius. 2004. Exemplar-based production of prosody: Evidence from segment and syllable durations. In *Speech Prosody 2004 (Nara, Japan)*, pages 459–462.

Katrin Schweitzer, Arndt Riester, Hans Kamp, and Grzegorz Dogil. 2008. Phonological and acoustic specification of information status - a semantic and phonetic analysis. Poster at "Experimental and Theoretical Advances in Prosody", Cornell University.

Kim Silverman, Mary Backman, John Pitrelli, Mari Ostendorf, Colin Wightman, Patti Price, Janet Pierrehumbert, and Julia Hirschberg. 1992. Tobi: A standard for Labeling English Prosody. In *Proceedings of ICSLP (Banff, Kanada)*, volume 2, pages 867–870, Banff, Canada.

Jacques Terken and Julia Hirschberg. 1994. Deaccentuation of words representing 'given' information: effects of persistence of grammatical function and surface position. *Language and Speech*, 37:125–145.

Jan P. H. van Santen and Bernd Möbius. 2000. A quantitative model of F0 generation and alignment. In A. Botinis, editor, *Intonation—Analysis, Modelling and Technology*, pages 269–288. Kluwer.

Michael Walsh, Hinrich Schütze, Bernd Möbius, and Antje Schweitzer. 2007. An exemplar-theoretic account of syllable frequency effects. In *Proceedings of ICPhS (Saarbrücken)*, pages 481–484.

Michael Walsh, Katrin Schweitzer, Bernd Möbius, and Hinrich Schütze. 2008. Examining pitch-accent variability from an exemplar-theoretic perspective. In *Proceedings of Interspeech 2008 (Brisbane)*.

Sandra P. Whiteside and Rosemary A. Varley. 1998. Dual-route phonetic encoding: Some acoustic evidence. In *Proceedings of ICSLP (Sydney)*, volume 7, pages 3155–3158.

George Yule. 1980. Intonation and Givenness in Spoken Discourse. *Studies in Language*, pages 271–286.