English Language and Linguistics 17.3: 537–563. © Cambridge University Press 2013
doi:10.1017/S1360674313000154

The role of prosodic structure in the formation of English blends

SABINE ARNDT-LAPPE and INGO PLAG
Heinrich-Heine-Universität Düsseldorf
(Received 22 February 2013; revised 4 June 2013)

This article investigates a variety of ways in which prosodic factors influence blend structure in English. Recent approaches no longer consider blends unpredictable, but the role of stress in blend formation has not been investigated in detail yet. This article addresses this problem, focusing on the role of stress in determining the switchpoint of the two bases in the blend, and on the question of what determines the stress pattern of the blend. We investigate these questions using experimentally derived forms, coined by native speakers on the basis of carefully controlled word pairs as stimuli. The results demonstrate that the length of the blend, the location of the switchpoint, and the stress of the blend are crucially determined by stress properties of the two base words of the blend, above all by those of the second word. At a theoretical level, the most important single finding is that preservation of the stress of the second word may happen independently of preservation of segmental material of the stressed syllable (e.g. *préstitant* from *prestígious* + *dóminant*). In contrast to stress, and contrary to earlier claims, syllabic constituency is shown to be of minor importance for switchpoint location. The theoretical implications of these findings are discussed. On a methodological level, our results show that experimentally elicited blends constitute a valid and highly useful resource for research on blend structure.

1 Introduction[1]

English has a number of morphological categories where the shape of the complex words belonging to that category is crucially determined with reference to syllabic or metrical structure. Examples of such 'prosodic morphology' (e.g. Plag 2003, Lappe 2007 for English; McCarthy & Prince 1993, 1998 for the general framework) are given in (1), with the diamond indicating a morphological relationship:

(1) (a) doc ◆ doctor, fax ◆ facsimile, gas ◆ gasoline, lab ◆ laboratory, mike ◆ microphone
(b) Bazza ◆ Barry, Dazza ◆ Darren, Gazza ◆ Gary / Gascoign, Hezza ◆ Heseltine
(c) brunch ◆ breakfast + lunch, buffeteria ◆ buffet + cafeteria, dramastic ◆ dramatic + drastic, britcom ◆ British + sitcom, prezactly ◆ precisely + exactly

The examples in (1c) are known under the label of 'blend'. Blends are derived on the basis of two (sometimes even more than two) bases. The first base (i.e. the left-hand one in writing) will be referred to as 'Word 1', the second as 'Word 2'. Being made

[1] The authors thank the audiences at several conferences for useful discussions of earlier versions: Manchester Phonology Meeting, May 2012; Data-rich Approaches to English Morphology, Wellington, July 2012; 35. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft, Postdam, March 2013. We are also very grateful for the feedback provided by the two anonymous *ELL* reviewers and the editor Bernd Kortmann. Special thanks to Manel Ben Abdallah for developing this project with us and for collecting the data, and to Ute Raffelsiefer for her help with the final coding of the data.

up of two or more bases, blends resemble compounds. In fact, Bauer, Lieber & Plag (2013: ch. 19) show that, apart from differences on the phonological level, blends show all crucial characteristics of compounds. The phonological difference to standard compounds is twofold. First, blends are syllabified and stressed like one phonological word, not like two, as compounds. Second, at least some of the phonological material is lost. This is further illustrated in (2).

(2)	**Word 1**		**Word 2**	**Blend**
	bréakfast	+	lúnch	brúnch
	fantástic	+	fábulous	fantábulous
	mótor	+	hotél	motél

With respect to the role of prosodic structure in English blend formation, two issues arise. The first is that, despite a number of pertinent observations and claims in the literature, the role of prosodic categories as determinants of blend structure is as yet far from clear. Thus, recent corpus-based work (esp. Gries 2004a, 2004b, 2006, 2012, to be discussed in detail in section 2) has provided evidence that the prosodic structure (especially word length and stress pattern) of blends is highly similar to that of their base words (especially to Word 2; see also Bat-El & Cohen 2012 on stress), and has interpreted this fact as an indication that similarity and recognizability play a role in blend formation. For stress, however, it is not clear how the observed similarity relations translate into predictive phonological restrictions that determine blend structure. For length, it is not clear whether the observed patterns are genuine length effects or epiphenomenal of other effects, most notably those of metrical structure preservation. Another prosodic category that has been claimed to act as a restriction on blend structure, but that has not yet been tested in detail empirically, is syllabic constituency. Kubozono (1990) and Bauer (2012), for example, propose that the switchpoint, i.e. the point where the two bases meet in the blend, occurs exclusively at syllable constituent boundaries (for example, between onset and nucleus, as in *smog* ◆ *smoke* + *fog*, or *brainiac* ◆ *brain* + *maniac*).

The second issue that arises with respect to the role of prosodic structure in blend formation is the question of how predictable blend structure is in general, and how predictability is related to the variation that we find in the data. In many earlier treatments of blends, these formations have been judged to be largely unpredictable (e.g. Marchand 1969; Bauer 1983: 225; Cannon 1986: 744). More recent approaches no longer consider blends unpredictable, and pertinent studies have found interesting restrictions at work that allow for the formulation of a number of interesting generalizations (e.g. Bat-El & Cohen 2012; Bat-El 2006; Bauer 2012; Gries 2004a, 2004b, 2012; Kubozono 1990).

A major problem for an assessment of variability in blend formation is, however, that all previous studies are based on (and generalize over) existing, mostly lexicalized, blends. This means that the variability problem has been investigated always across types, not within types, so that we never see the full range of structural options that are available for a single pair of base words. This problem is tacitly acknowledged by

Gries (2004b), who argues that *breakfunch* would be a better or more likely coinage than the existing blend *brunch*.

A serious methodological disadvantage of research being conducted solely on the basis of existing blends is that certain effects are very hard to measure because each word brings in its own structure, so that the conditions for the operation of a particular restriction cannot be held constant. This problem is aggravated by the fact that, as shown in detail by Gries (2004a, 2004b, 2006, 2012), blending especially occurs among structurally highly similar base words, so that any corpus-based data set is going to be highly biased in terms of the structural configurations to be encountered among base words. Configurations which are rare, but provide an interesting insight into the workings of structural determinants, are therefore difficult to investigate on the basis of existing types.

Finally, it is unclear how the fact that a blend is lexicalized affects the question of how well its structure reflects productive mechanisms. Thus, it is a common assumption in corpus-based work on productivity that productive mechanisms are best investigated in rare formations rather than in frequent formations (e.g. Baayen 1992, 1993; Plag 1999). Frequency of occurrence has, however, never been an issue in corpus-based studies of lexical blending.

In the present article, we will use an experimental methodology to investigate the role of prosodic structure in English blend formation. This methodology involves a production experiment in which native speakers of English were asked to coin blends on the basis of carefully controlled word pairs as stimuli. In the analysis of the data, we will focus on two questions:

- How is the switchpoint between the two bases determined?
- What determines the stress pattern of the blend?

Our contribution to research on blending in English is twofold. On a methodological level, our research will complement earlier, corpus-based studies. We will see that experimentally elicited blends constitute a valid resource for research on blend structure, which allows us to study both within-type variability and structural configurations that are underrepresented in lexicalized data in more detail than other methodologies. On the level of contents, our study will provide new insights into the way prosodic categories shape the structure of English blends. Specifically, categories of syllabic constituency seem to be less relevant than hitherto assumed. By contrast, metrical categories will be shown to be more relevant than noted in the literature, shaping not only the stress pattern of blends as compared to their source words, but also having a crucial influence on the overall structure of the blend and the location of switchpoints. These empirical findings have important implications for theoretically oriented models of blend formation.

The article is structured as follows. In section 2 we review in more detail the problems in the description of blends and their structure, and develop our research questions. Section 3 presents our methodology. This is followed by section 4, in which we present the results of our analysis. Section 5 provides a summary and a final discussion.

2 The role of prosodic structure in English blends

Morphologically complex words are semantically and formally related to their bases. Leaving semantic issues aside, it is safe to say that formal relatedness is often not straightforward. Thus, we find stress shifts, base allomorphy, deletion of segments or other kinds of phonological processes that impinge on how faithful the complex word is to the phonological structure of its base(s). With blends, this is an especially intricate problem, which can be broken up into at least three subproblems, all of which seem to show variability. First there is the problem of anchoring (see, for example, Lappe 2007 for discussion): from which point onwards in the string, and in which direction, does material from a base survive? Note that the notion of switchpoint introduced above refers to the anchoring points of the two bases from the perspective of the blend. The point that constitutes the switchpoint for the blend is the anchoring point for the bases. The second problem is length, i.e. how much of the bases survives. The third problem is the metrical structure, i.e. how stress is assigned. We will discuss each in turn.

2.1 Anchoring

We have phrased the anchoring problem by focusing on what material survives from the base. This perspective is relatively recent and many earlier studies such as Bauer (1983) or Cannon (1986) have looked at what is deleted instead. However, as shown, for example, in Lappe (2007) and Alber & Arndt-Lappe (2012), an output-oriented approach to prosodic morphology, which uses constraints on survival rather than on deletion, is empirically and theoretically more adequate. We will therefore adopt the same kind of approach in this article.

2.1.1 Left and right anchoring

It has often been observed that with regard to anchoring, there may be a lot of variation. There are two main patterns for the coinage of compounds that concomitantly lose phonological material in this process, given in (3), with the pattern in (3b) being rather marginal (see Plag 2003: ch. 5.2.2; Bauer, Lieber & Plag 2013: ch. 19, section 4.8).

(3) (a) AB + CD → AD *br*eakfast + l*unch* → *brunch*
    (b) AB + CD → AC *mo*dulator + *dem*odulator → *modem*

In one pattern (3a), the initial part of Word 1 and the final part of Word 2 survive, while in the other pattern (3b) the initial parts of both bases survive. In the present study, only 6.5 per cent of our data set of 1,357 formations could not be subsumed under the AD template. This small proportion is in line with what previous studies have found (Kubozono 1990; Gries 2006). AD is clearly the much preferred pattern. Furthermore, as demonstrated by Gries (2006), AC formations and AD formations follow different structural requirements. AC formations are therefore often treated as a pattern distinct from blending, referred to as ‘clipped compounds’ (Bauer, Lieber & Plag 2013) or ‘complex clippings’ (Gries 2006). Clipped compounds systematically preserve much less material than AD formations, and overlaps of segmental material are less common in clipped compounds. Finally, AD blends and clipped compounds differ with regard to their base words. While bases of AD blends tend to be orthographically and phonologically highly similar to each other, clipped compound bases are significantly less similar to each other.

We therefore adopt a definition of blend according to which blends are two-constituent compounds in which at least one constituent has lost some phonological material, and in which the left, or initial, part of Word 1 and the right, or final, part of Word 2 survives. We say ‘at least one constituent’, as there are cases which can be analysed in such a way that one of the bases survives completely (e.g. *painstation* ◆ *pain* + *playstation*, *wreader* ◆ *writer* + *reader*). Such forms can be accommodated by the AD template by stating that the respective ‘part’ may be as large as the whole word.

As for variation, the anchoring points (or, from the perspective of the blend, the switchpoints) are an obvious source of variation, with partial vs full preservation of a base as one specific instance. There are, however, some more specific claims about the determinants of the location of the switchpoints, to which we now turn.

2.1.2 The location of switchpoints

As already mentioned above, Kubozono (1990) and Bauer (2012) hold that switchpoints occur only at syllable-constituent boundaries, or at boundaries of syllabic constituents. In particular, Bauer (2012) claims that in monosyllabic blends the switchpoint is between the onset of Word 1 and the rhyme of Word 2 (*brunch* and *smoke* are examples exhibiting the expected pattern). The general idea that syllabic constituency has some relevance in the determination of switchpoint location also receives empirical support from Gries (2012), who shows that in intentional blends switchpoint location tends to be at syllable-constituent boundaries. However, his study also documents instances where this generalization does not hold. The status of these documented cases is, however, unclear.

Gries (2006) comes up with a non-structural factor, which he calls the ‘selection point’. This point is close to the ‘recognition point’, i.e. ‘the point at which a majority of speakers (e.g., 85%) can recognize [a word] *W* with a high probability (e.g., 80%) when presented with parts of *W*’ (see Gries 2006: 539–45 for full discussion). In Gries’s statistical models the selection points of Word 1 and Word 2 are highly significant predictors of switchpoint placement. It is unclear, however, how far this psycholinguistic effect coincides with syllabic constituency.

In this article we not only test the role of syllable structure in determining blend switchpoints, but we also propose and test a new hypothesis. This hypothesis states that stress plays an important role in the determination of switchpoints. In particular, we investigate whether the switchpoints depend on the metrical structure of Word 1, Word 2, or the blend.

2.2 Length

One can find a number of proposals in the literature concerning the length of blends, and how this length is related to the length of the bases. In general, length is most often discussed in terms of number of syllables. There is general agreement that especially the length of Word 2 is an important correlate of the length of the blend, but less agreement about (a) the exact role of Word 2, and (b) the potential interaction of Word 2 effects with those of Word 1. With respect to (a), we find conflicting hypotheses in the literature. Whereas Bat-El (2006) claims that the length of the blend is identical to the length of Word 2, Bauer (2012) offers the more moderate hypothesis that blends may not be longer than Word 2. Kubozono (1990), by contrast, claims that blends show a tendency to be minimally as long as Word 2. Gries's (2004a) study supports Kubozono's predictions. Both Gries (2004a) and Kubozono (1990), however, make it clear that the patterns observed in the literature with respect to blend length are tendencies. It is not clear whether and in how far violations of these tendencies are systematic.

Apart from the length of Word 2, also Word 1 length has been identified as a potential determinant of blend length. Most notably, Cannon (1986: 741) finds that blend length is identical to the length of the longer one of the two bases (see also Bat-El 2006: 67). This is similar to what was recently proposed for German blends (Costa 2008).

In the present study we will systematically compare blend length to both Word 2 and Word 1 length. The results of our analysis of metrical determinants of switchpoints, however, clearly call into question the status of length of the source words as a predictor of blend structure in its own right. Instead, we suggest that it may be an epiphenomenon of metrical structure preservation. This interpretation is fully in line with the assumption that recognizability of the source words is a crucial factor in blend formation (see Gries 2004a, 2004b, 2006, 2012).

2.3 Stress

There is not much discussion in the literature on English blends concerning stress. Cannon (1986: 741, 746) proposes that the stress of the longer word also determines the primary stress of the blend. Bat-El (1996), Fischer (1998) and Bauer (2012) have proposed that blend stress corresponds to Word 2 stress. In other words, if Word 2 is stressed on the final syllable, so will the blend be; if Word 2 is stressed on the penult, so will the blend be, etc. A similar proposal has been made for Spanish blends by Piñeros (2004), who says that blends in this language replicate the prosodic structure of Word 2, i.e. they show the same number of syllables and the same stress pattern as Word 2. An example would be *locómbia*, based on Word 1 *lóco* and Word 2 *Colómbia*.

The hypothesis that blend stress correlates with Word 2 stress in English seems to receive support from Gries's (2004a) study, which shows that blend stress tends to correspond to Word 2 stress. A methodological problem encountered in empirical studies of blend stress like Gries's is, however, how to measure correspondence of stress between participant words. Thus, for example, Gries's coding of stress patterns is intricately linked to the length and segmental material in the blend. This has the consequence that, for example, the blend *mírthquake* is coded as having the same stress pattern as Word 2 (*éarthquake*), but not as Word 1 (*mírth*). Unfortunately, such a coding does not allow any conclusions about the interaction of stress assignment and stress preservation. In the above example Gries's coding misses the very interesting fact that the penultimate main stress on the blend (*mírth.quake*) preserves both the main stress of *mirth* and the stress pattern of Word 2 (*éarthquake*). We will use a coding that is independent of word length to systematically test generalizations concerning stress assignment and stress preservation effects in blend formation.

3 Methodology

3.1 Experimental setup

In a production experiment we asked 29 native speakers of Irish English (all students at the University of Colraine, Ireland) to create blends on the basis of two words that were presented to them (Ben Abdallah 2008). The experiment consisted of a written and an oral part. In the written task, the participants were presented with a questionnaire that presented 60 pairs of words in a carrier sentence, and the participants were asked to create one word from the two source words and write it down. They were then asked to pronounce the word they had created. The oral production was recorded and later transcribed phonemically. The written data were collected for two reasons: to give the participants sufficient time to find an acceptable blend, and as a backup to ensure a correct interpretation of the audio data.

The stimuli consisted of word pairs that had the same word class and could together be interpreted as a coordinative compound. This was done to make the task as natural as possible, since attested blends tend to be of the coordinative type (even though blends can in principle exhibit the same range of semantic relationships as non-blend compounds; see Bauer, Lieber & Plag 2013: ch. 13). The bases were chosen in such a way that they systematically varied in terms of number of syllables and stress placement. Note that, due to the numerical explosion of combinations of properties (see the next subsection for an overview of the many variables at play), it was not possible to come up with a set of items that was fully balanced with regard to all sorts of constellations of base and blend properties. Thus, the subsets of forms featuring a particular combination of particular properties in bases and blend vary in size. This is, however, not detrimental to the purposes of this study. Example (4) lists some stimuli together with the blends as created by the participants. The complete list of stimuli can be found in appendix 1.

(4)	bar	+	restaurant	◆	bártaurant, báraurant, béstaurant
	jacket	+	coat	◆	joat, jácoat, jacóat
	pan	+	pot	◆	pánot, pat
	singer	+	actor	◆	singáctor, sáctor, síngtor

After the elimination of unusable data (missing responses, illegible responses or responses that were not clearly audible), we obtained 1,357 formations produced on the basis of 107 word pairs. The discrepancy between the 60 stimuli pairs and the 107 word pairs in the resulting data set arose from the fact that sometimes the participants chose to ignore the sequence of the two bases as given in the questionnaire and switched

Table 1. *Linguistic categories coded*

Variable	Values
surviving part of Word 1	initial, medial, final, complete
surviving part of Word 2	initial, medial, final, complete
medial overlap	yes, no
length	number of syllables (of Word 1, of Word 2, of the blend)
stress from right	number of syllables from the right (of Word 1, of Word 2, of the blend)
stress from left	number of syllables from the left (of Word 1, of Word 2, of the blend)
switchpoint Word 1	number of syllables from the left
switchpoint Word 2	number of syllables from the right
location of switchpoint	at syllable constituent boundary: yes, no
constituent boundary of switchpoint	onset, nucleus, coda

Word 1 and Word 2, thus producing a new word pair and a new blend. All pairs entered the analysis on the grounds that we would not expect blends produced on the basis of a switched order of bases to be subject to different structural mechanisms. This assumption was supported by an analysis of the switched pairs in terms of the pertinent structural categories, which did not reveal any clearly distinct patterns that could have motivated the switches. Finally, recall that all formations elicited in the experiment had a coordinative meaning. We therefore do not expect the reversal to be detrimental.

3.2 Categories coded

The two bases and the blends were coded with regard to a range of linguistic properties. All codings were made on the basis of the pronunciations, not on the basis of the forms as written down in the questionnaire. The categories that are relevant for this article are listed in table 1.

To analyse the overall structural pattern we coded which part of the respective bases survived in the blend, and whether there is overlap of Word 1 and Word 2 material in the blend ('medial overlap', as in *jacoat*, in which the two /k/'s of the bases overlap). We measured the length of the blend, of Word 1, and of Word 2 in number of syllables. We also counted the number of syllables to code the stress patterns of Word 1, Word 2 and the blend. This was coded in two ways, counting from the left and counting from the right. The location of the switchpoint was coded in different ways: we noted in which syllable of the bases the switchpoint occurred, counting from the left for Word 1, and from the right for Word 2. We also coded whether the switchpoint was at a syllable-constituent boundary, and if so, at which kind of boundary. Table 2 illustrates the coding for a token of the pair *prestigious* + *dominant* with the resulting blend /prəs.ˈtɪ.dʒɪ.nənt/.

Table 2. *Illustrative coding for Word 1:* prestigious*, Word 2:* dominant*, blend:* /prəs.ˈtɪ.dʒɪ.nənt/

Variable	Values
surviving part of Word 1	initial
surviving part of Word 2	final
medial overlap	no
length Word 1	3
length Word 2	3
length blend	4
stress from right Word 1	2
stress from left Word 1	2
stress from right Word 2	3
stress from left Word 2	1
stress from right blend	3
stress from left blend	2
switchpoint Word 1	3
switchpoint Word 2	2
location of switchpoint	constituent boundary
constituent boundary of switchpoint Word 1	onset
constituent boundary of switchpoint Word 2	nucleus

As one can see, the codings were done in such a way that theoretical biases were avoided as far as possible to allow for a description of the facts in a way that is as theory-neutral as possible.

3.3 Further methodological considerations

The methodology used here may raise some methodological concerns, which we would like to address in this section. The first is that, given the design of the study, the formations obtained here are forced formations, for which it is not quite clear how many of them would actually be realistically attested as blends in the language. An example is the pair of source words *pan* (Word 1) and *pot* (Word 2), which, in the majority of cases, made our informants create the blends *pánot* or *panót*. Whereas anecdotal evidence from native-speaker intuitions may suggest that *pan* and *pot* are odd bases for blends, the complexity of the many factors involved makes it difficult to estimate on principled grounds exactly which bases would lead to a null parse. This objection would therefore make it generally difficult to elicit blends in an experimental setup. What makes us optimistic in terms of the quality of our data as quite realistic blends is that, in some instances, speakers did refuse to form blends from the base words given, but that these instances were few in number and did not show a systematic pattern. We interpret this as an indication that our informants may differ in terms of their competence with respect to the set task, and that the difference between suitable bases for blending and unsuitable bases is likely to be a gradual one. Therefore, we have no principled basis for making a sharp distinction between possible and impossible bases for blending.

Another, somewhat related, issue is that the source words in our experiment differ substantially from the source words that are attested for naturally occurring blends. In particular, Gries (esp. 2012; see also 2004a, 2004b, 2006) shows that lexicalized blends (termed 'intentional blends' by Gries) are characterized by a relatively high degree of similarity between the two source words. This pertains to both their segmental makeup and their stress pattern. For the study of phonological determinants of blend structure, however, this means that some configurations (in terms of segmental and prosodic structure) are overrepresented in lexicalized data whereas others are extremely rare, so that an analysis can sometimes contribute little to an assessment of current hypotheses about determinants of blend structure. This becomes especially clear when we want to study prosodic determinants of blend structure. Given that, in spite of some variation, word stress assignment in English is far from random, it is to be expected that we will not find too much variation in stress patterns of source words for lexicalized blends. Gries (2012: 154) confirms exactly that expectation: 'when the source words of intentional blends have the same number of syllables, then there is also a significant tendency for them to have the same stress pattern'.

In order to study stress assignment in blends, then, lexicalized data are not the most informative type of data because what we need to see is exactly the rare cases where source words exhibit different stress patterns. Also, the blends investigated here differ from existing blends in terms of the lengths of the source words. Whereas Gries (2012: section 3.1) shows that the source words for lexicalized blends tend to be of unequal length (with Word 2 being significantly longer than Word 1), stimuli in our experiments are more balanced with respect to different length configurations, which allows us to better test source word length as a determinant of blend length.

4 Results

4.1 Overview of the data

As already mentioned in section 2.1.1, applying the above definition of blend to our data set, we end up with 1,269 blends, i.e. AD formations. Of these, 304 blends have one constituent that has not lost any material (24 per cent). In roughly two-thirds of such cases (68 per cent), it is Word 2 that is fully preserved. Again this is in line with previous observations in the literature (e.g. Bauer 1983: 235), and it is also a reflection of the general tendency of blends to preserve more material from Word 2 than from Word 1 (see also Gries 2004b). Example (5) illustrates the variability of full preservation with the pair *publisher* + *editor* from our data set. Four different blends were produced, two of which preserve Word 2 completely and two of which don't.

(5) publisher + editor ◆ publéditor, pubéditor, púbitor, públitor

In 16.5 per cent of the 1,269 blends there is a medial overlap, a rather common phenomenon (see, for example, Bauer 1983: 235; Cannon 1986: 741). In the vast

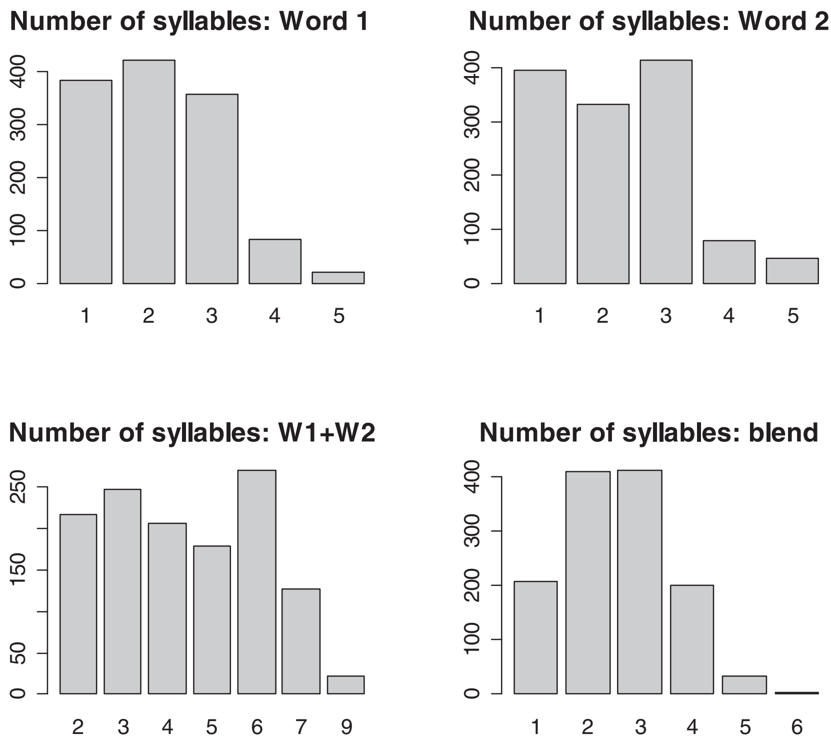

Figure 1. Length of bases and blends

majority of our cases the overlap concerns only one segment, which can be a vowel (*journalist* + *moderator* ◆ *journ**a**rator*) or a consonant (e.g. *scanner* + *printer* ◆ *sca**n**ter*).

With regard to the length of the blends produced by our participants, we find the distributions as shown in figure 1. As shown in the two upper plots, there is a rather even distribution of monosyllabic, disyllabic and trisyllabic bases in our stimuli set. These words are also fairly evenly distributed over the word pairs, as we can see from the lower left plot, which gives the frequency distribution of the sums of the syllable counts of Word 1 and Word 2. If we compare this plot with the one giving the length distribution of the resulting blends (lower right plot), we can see that there is a general preference for disyllabic and trisyllabic blends. This suggests that the longer the base word the more syllables tend to get lost. The correlation is indeed highly significant ($rho = 0.83$, $p < 2.2\text{e-}16$, Spearman).

With regard to the question of what role Word 1 and Word 2 play in determining the length of the blend, our data conform to the hypothesis according to which blends are maximally as long as their longest source word (Cannon 1986; Bat-El 2006: 67; see also Kubozono 1990 or Gries 2004a for similar findings). Only 9 per cent of all blends in our data that have bases of different lengths do not behave according to this hypothesis.

Among the blends with bases of the same length the error rate with this hypothesis is 23 per cent, which is still better than with any of the competing hypotheses.

To summarize, our experimentally derived blends seem to be very similar to existing blends in terms of their structural properties. This means also that, in spite of the artificiality of the experimental setup, the speakers came up with coinages that match coinages that were produced in more natural settings. This supports the idea that an experimental approach to blends is viable and yields valid data.

We may now turn to the phenomena that are in the focus of this study, switchpoints and stress.

4.2 Switchpoints

4.2.1 Constituent boundaries

Previous studies (e.g. Kubozono 1990; Bauer 2012) have proposed that switchpoints occur exclusively at constituent boundaries, for example between onset and coda, as in *smog* (*smoke* + *fog*), or *brainiac* (*brain* + *maniac*). In our data set, switchpoints are indeed mostly at syllabic constituent boundaries, i.e. after or before an onset, nucleus or coda, but not necessarily so (see Gries 2012). This somewhat unexpected behaviour happens in less than 4 per cent of all blends, which may look like a rather exceptional phenomenon that can easily be dismissed. However, this low figure is due to the fact that there is an uneven distribution of consonant clusters in the bases. A closer inspection of the data is therefore necessary, and indeed rewarding.

Let us look at the onset–nucleus boundary and the word pairs in which Word 1, or both Word 1 and Word 2, begin with a complex onset. In such cases, there are in principle quite a number of switchpoints possible, for example, after the onset of Word 1, inside the onset of Word 2, or after the onset of Word 2. If, contra Kubozono and Bauer, no special restrictions are at work in blends with regard to the placement of switchpoints inside onsets, we would expect that switchpoints may occur inside onsets, as long as the newly created onset does not violate the phonotactic constraints of English. For example, *black* + *white* may not become *bwhite* because of the illegal onset, but *blue* + *green* may become either *breen* or *bleen*, or *green* + *blue* may become *grue* or *glue*. Similarly, *scanner* + *printer* could become *s*[k]*inter*, *scrinter*, or *sprinter*. This prediction is in fact borne out by the data. All 60 cases (tokens) in our data where the new combination would be illegal show a switchpoint at the onset–nucleus boundary, and not inside the onset. In contrast, of the 40 cases (tokens) where the new combination of sounds would result in a legal onset cluster, we find a preference (of 70 per cent, N = 28) for placing the switchpoint at the boundary, but also 30 per cent of switchpoints inside the onset (N = 12). The quantitative difference resulting from the categorical behaviour of the former group of blends and the variable behaviour of the latter group of blends is statistically highly significant (p = 5.4e-06, Fisher test, see also table A1 of appendix 2).

If we look at the types instead of the tokens we find a similar picture. All combinations that have the option of a switchpoint within a legal cluster do in fact show a variable placement of the switchpoint. In other words, with such types there is always at least one participant for each type who coins a blend with a switchpoint inside the onset cluster. These distributions of switchpoints inside onset clusters vs before or after onset clusters strongly suggest that there is no general restriction of switchpoints to syllable constituent boundaries. Instead, complex onsets in source words may variably split, and the parts may be variably recombined or deleted, as long as the resulting combinations of sounds are phonotactically legal. This is illustrated in (6), with '|' indicating the switchpoint.

(6) (a) *Split and recombination*

blue + green	→	bl|een, b|reen
scanner + printer	→	sc|inter, sc|rinter, s|printer

(b) *Split and deletion*

publisher + editor	→	publ|editor, pub|editor
employer +director	→	empl|irector, emp|irector
lecturer + tutor	→	/lek|jutəɹ/

Based on the across-type and within-type variation found in our data, Kubozono's and Bauer's hypothesis that switchpoints cannot occur inside onsets must be rejected. We do not find switchpoints inside complex nuclei or complex codas, however. Further research is necessary to establish whether this patterning stands the test of broader empirical investigations and, if so, whether it results from independent restrictions. Let us now turn to the role of stress in determining the switchpoint.

4.2.2 The role of stress in the determination of the switchpoint

The literature is silent about whether stress plays a role in determining the switchpoint. On the basis of several properties of blends and independent considerations concerning the general treatment of prominent material in morphology, one can, however, set up a hypothesis about the relationship between stress and switchpoints.

We know from other morphological categories, and especially from those belonging to the realm of prosodic morphology, that prominent material from the bases tends to be preserved more faithfully than non-prominent material (e.g. Benua 2000; Lappe 2007; Alber & Arndt-Lappe 2012). Thus, initial and stressed material has a higher chance of preservation than unstressed material, or than material that is located further towards the end of the base. These facts about the phonological relation between morphologically related words can be seen as motivated by psycholinguistic processes. Thus, what tends to be faithfully preserved in morphological alternations largely corresponds to what has been found to be important in lexical processing (see, for example, Alber 2001; Lappe 2007).

Blends preserve the initial part of Word 1 and the final part of Word 2. Now given that initial material is prominent, and Word 2 loses its initial material, Word 2 must somehow compensate for the loss of its initial material in order to remain recoverable. One mechanism is to preserve as much non-initial material as possible, and indeed it has been observed that on average Word 1 loses more material than Word 2 (e.g. Gries 2006). Another mechanism might be to preserve the stressed syllable of Word 2. Thus

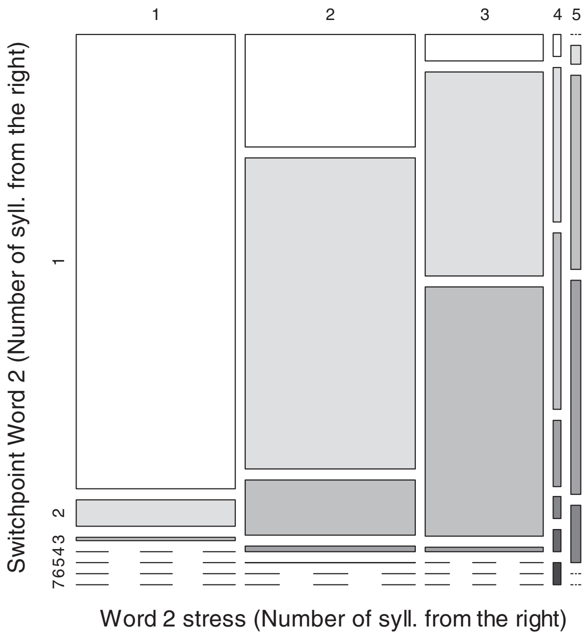

Figure 2. Switchpoint by Word 2 stress

we could hypothesize that the switchpoint is no further to the left than the stressed syllable of Word 2. Roughly speaking, it would be enough for Word 2 to preserve its stressed syllable and everything to the right of it.

A quantitative analysis of our data strongly supports this hypothesis: figure 2 gives a mosaic plot for the distribution of switchpoints by Word 2 stress for all blends that have more than one syllable. Mosaic plots represent the number of observations in each subset of the data as the corresponding proportion of the plotting area. A large area thus represents many observations with a particular combination of properties, a small area represents a proportionally smaller number of observations The x-axis gives the stressed syllable of Word 2, counted from the right, and the y-axis gives us the location of the switchpoint in Word 2, counted in syllables from the right. The table underlying the plot shown in figure 2 is given as table A2 in appendix 2.

There is a clear relation between the two variables ($p < 0.001$, Fisher test with simulated p-values, based on 2000 replicates). We can see that for the majority of Word 2s with final stress (value of Word 2 stress = 1) the switchpoint is also in the final syllable of Word 2 in the blend (leftmost white area). There are only a few violating forms (in the two areas below). Basically the same kind of distribution holds if Word 2 has penultimate stress, or antepenultimate stress. Overall, the error rate for our hypothesis is only 9.1 per cent (N = 1,263, six observations were lost due to missing values).

The switchpoint can be at the left boundary of the stressed syllable of Word 2, within or after the onset of the stressed syllable of Word 2, as illustrated in (7).

(7) scánner + prínter → sc|rínter, sc|ínter
gigántic + enórmous → gi|nórmous
Brítish + Índian: → Br|índian
emplóyer + diréctor → empl|éctor

If we look at the types instead of the tokens, we find further evidence for the reality of the proposed restriction. All word pairs that yield a blend violating the generalization also yield a blend variant that does not violate the generalization. See the examples in (8):

(8) díctionary + thesáurus → dic|thesáurus (*violator*)
díctionary + thesáurus → dic|sáurus (*non-violator*)

Another interesting subpattern can be discerned for those pairs in which the number of violators is larger than the number of blends that conform to the hypothesis. With this subset we find that the two bases end in the same sequence of sounds (*consultátion* + *interpretátion*). In all such cases there is a preference for blends that preserve material beyond the stressed syllable of Word 2. See (9), in which the parts of Word 2 that are to the left of the main-stressed syllable are in bold print.

(9) consultátion + interpretátion → consul|**pre**tátion, con|**pre**tátion

The reversal of preferences for this particular subset of pairs indicates that deletion and preservation are ultimately determined by the degree of recoverability of the two bases. In cases of final overlap more of Word 2 needs to be preserved in order to make it distinguishable from Word 1, with which Word 2 shares the same material at the end. From this observation one could venture the hypothesis that the larger that overlap the more probable a violation of our hypothesis would become. Future research will have to test this, as our data set does not provide enough suitable forms.

4.3 *Blend stress*

4.3.1 Results

Recall from section 2 that there are two competing views. One view focuses on size (Cannon 1986: 741, 746, the longer word provides the stress pattern of the blend), the other on position (Bat-El 1996; Fischer 1998; or Bauer 2012, Word 2 provides blend stress). Bat-El & Cohen (2012) advocate an intermediate position with both position and size playing a role, depending on very specific prosodic constellations of Word 1, Word 2, and the blend.

We will first test Cannon's prediction that the longer word provides the stress pattern of the blend. In our data set, 365 tokens feature a longer Word 2, while 288 have a longer Word 1. Those blends that have a longer Word 1 are the crucial test cases for Cannon's hypothesis, since for those blends that are based on a pair with a longer

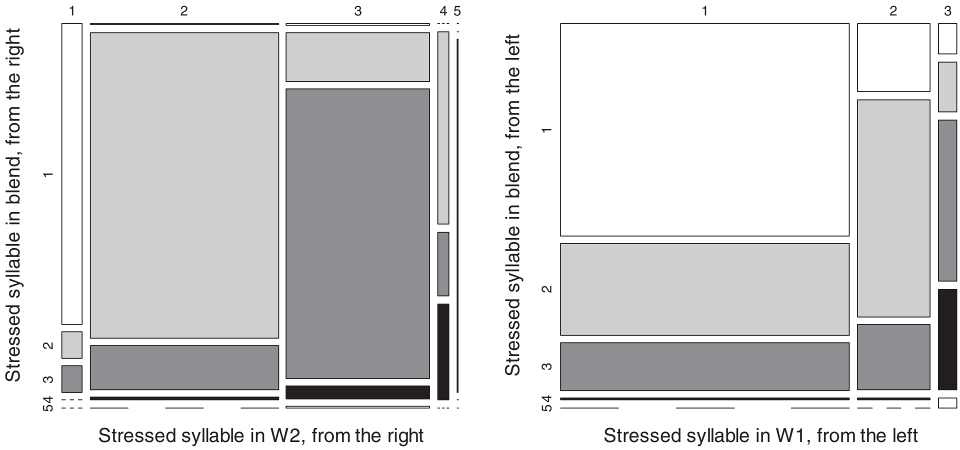

Figure 3. Blend stress by Word 2 stress and by Word 1 stress

Word 2, the size-based hypothesis makes the same prediction as the position-based hypothesis, namely that blend stress is the same as Word 2 stress. This prediction is borne out by the data, 82 per cent of the blends with a longer Word 2 stress the same syllable as Word 2.

In those blends that have a longer Word 1, however, only 53 per cent are stressed on the same syllable as Word 1, with stress counted from the left. This low rate drops further to 42 per cent if stress is counted from the right. This clearly refutes Cannon's hypothesis. A potential reason why Cannon comes up with this hypothesis in the first place is that in most existing blends Word 2 is the longer word anyway (for reasons discussed in the previous section).

Let us now turn to the position-based hypothesis. Figure 3 shows the relevant distribution of stresses in our data. The left panel shows the relation between Word 2 stress and blend stress, while the right panel shows the relation of Word 1 stress and blend stress. We measure the stress relation between the two bases and the blend by counting syllables from the base edges that survive in the blend. That is, for the relation of stresses between Word 2 and the blend we count syllables starting from the right edge, and for the relation between Word 1 and the blend we count syllables starting from the left word edge. The plots are based on a subset of the data in which all words involved, i.e. Word 1, Word 2 and the resulting blend, are polysyllabic to allow for maximal variation (N = 705). The two tables underlying the two plots are given as tables A3 and A4 in appendix 2.

Each plot in figure 3 shows a clear dependency of the two variables involved ($p < 0.001$ for both distributions, Fisher test with simulated p-value based on 2,000 replicates), but the effect of Word 2 is much more pronounced. Differences between the position of stress in the blend and the position of stress in Word 2 appear in only 17 per cent of the cases, while the prediction of blend stress on the basis of Word 1 stress is much less successful (40 per cent failure of Word 1 stress preservation, the difference is very highly significant, $p = 1.994\text{e-}12$, *chi-square* = 49.4899, *df* = 1). In sum, blend stress is chiefly determined by Word 2 stress.

If blend stress is dependent on Word 2 metrical structure built up from the right, what happens pretonically? If we inspect both blend stress and Word 2 stress by counting from the left, there is much more variation than in the analysis that looked at stress from the right (as related to Word 2 stress). The initial part of blends very often does not show the prosodic structure of Word 2 (34 per cent violators). The difference between this result and the one for stress being counted from the right (with only 17 per cent violations) is highly significant (*chi-square* = 29.8912, *df* = 1, $p = 4.57\text{e-}08$), which means that we find significantly more variation pretonically. This increase in variability can be interpreted as resulting from a conflict between the preservation of material from Word 1 and the preservation of the Word 2 stress pattern. The more material we preserve of Word 1, the more we run the danger of losing faithful preservation of Word 2 prosody since more material from Word 1 increases the chances of also bringing in stressed material additional to that of Word 2.

The dominant role of Word 2 stress in stress assignment to the blend raises the question of the relation between segmental and metrical faithfulness. Many of the words in second position are stressed on the initial syllable (e.g. *móderator*), and at the same time this initial syllable often does not survive in the blend (e.g. *jóurnarator* ◆ *jóurnalist* + *móderator*). This leads to the obvious and interesting question whether the Word 2 stress pattern survives in the blend even if the segmental content of the stressed syllable of W2 does not survive. This can be answered in the affirmative. There are 286 cases in our data where the vocalic nucleus of the stressed syllable of Word 2 is not preserved in the blend, but the blend still shows the same stress pattern as Word 2 in 67 per cent of these cases. In many of these cases, however, Word 1 stress and Word 2 stress are identical, as for example, in the case of *públisher* + *éditor* (◆ *públitor*, for example) or *módify* + *símplify* (◆ *módlify*, for example). It is therefore instructive to look at those cases where Word 1 stress and Word 2 stress (counted from the right) differ, as in *prestígious* + *dóminant* (◆ *préstinant*, for example), or *jóurnalist* + *móderator* (◆ *jóurnarator*, for example). This subset has 122 items, of which 52 per cent of the blends are stressed like Word 2, 30 per cent like Word 1, and 17 per cent do not show either stress pattern. This shows that the stress pattern of Word 2 has a strong tendency to survive, independent of the survival of the segmental material of the stressed syllable.

Interestingly, when counting from the left, we also see a tendency towards stress preservation of Word 1. Thus, in 62 per cent of the 122 items, the first syllable of Word 1 is stressed, and the first syllable of the blend is stressed. This effect comes about because by definition the first syllable of Word 1 survives, and if this syllable is stressed, its stress can survive easily in those cases where the first syllable of Word 1 ends up in a position of the blend which corresponds to the stressed position of Word 2. Cases in point are, for example, *jóurnalist* + *móderator* ◆ *jóurnarator*, or *wríter* + *photógrapher* ◆ *wrítgrapher*. Notably, in these cases there is no conflict between stress preservation from the left, i.e. Word 1, and stress preservation from the right, i.e. from Word 2.

Thus, the empirically most adequate analysis of blend stress is one in which Word 2 provides the metrical structure for the blend, irrespective of where the segmental material is taken from, Word 1 or Word 2. At a more abstract level, we can say that segmental and prosodic faithfulness work independently of each other.

Let us return to those cases where, counting from the right edge, stress assignment is not based on Word 2 prosody. These cases are interesting, as they show two clear patterns. These blends either show the stress pattern of Word 1, or have a stress pattern expected for monomorphemic words. For illustration, consider the combination of *journalist* + *moderator*, for which we have 22 observations. This combination is the only one with a quadrusyllabic Word 2, which is the reason why the respective column of figure 3, left panel, is quite narrow. Examples of blends formed for this combination by our subjects are given in (10), together with the corresponding frequencies of the patterns.[2]

(10) (a) *Word 2 stress preservation*
jóur.na.ra.tor (*N* = *6*)
(b) *Monomorphemic stress patterns*
jour.na.rá.tor (*N* = *11, stress on heavy penult*)
jour.ná.l[ə].ter (*N* = *4, light penult, stress on antepenult*)

Further support for the analysis gleaned from the pair *journalist* + *moderator* comes from pairs where Word 2 is disyllabic. We find within-type variation and most violators preserve W1 stress (e.g. *éx.port* + *ím.port* → *éx.im.port*). If Word 1 is longer than Word 2, Word 1 stress preservation is actually preferred, as in *sé.mi.nar* + *lécture* → *sé.mi.ture* or *lécturer* + *tú.tor* → *léc.tu.tor*.[3]

To summarize, we have found robust evidence for the generalization that Word 2 prosody largely governs stress assignment to blends. Exceptions to this pattern do occur and they are systematic in the sense that they either involve the preservation of Word 1 prosody, or show the treatment of blends as monomorphemic words by the subjects. Under which exact circumstances Word 1 prosody prevails over Word 2 prosody remains to be studied in more detail.

4.3.2 Discussion

Our results show an overwhelming tendency for blends to preserve the prosodic structure of Word 2. This not only refutes a strictly size-based view of blend stress as advocated by Cannon (1986), it also runs counter to some of the claims in Bat-El & Cohen (2012). Using an optimality-theoretic approach, these authors present

[2] There is only one exceptional form with this pair, not given in (10). This exception concerns the blend *jour.náler*, whose structure also deviates in terms of switchpoint location.

[3] The tokens in the group of three-syllable Word 2s that violate faithfulness to Word 2 prosody do not lend themselves to a systematic analysis, as they consist of erroneous forms or do not allow for an uncontroversial counting of the number of syllables.

the hitherto most detailed analysis of blend stress, and it is therefore interesting and informative to discuss their claims in the light of our new empirical findings.

Bat-El & Cohen (2012) also propose that Word 2 prosody has a strong influence on blend stress, and they also note the two exceptional patterns that we found in our data. In their attempt to provide a principled account of these patterns, the authors distinguish numerous different constellations, depending on the sizes of the bases and the blend, and depending on which stressed base syllables survive in the blend. Unfortunately, the classification of their data does not correspond to our coding of the experimental data so that a direct general comparison of the pertinent subsets of data is not possible, but it is nevertheless possible to discuss some important aspects.[4]

For example, in our data we find a clear preponderance of Word 1 stress preservation only with polysyllabic bases where Word 1 is longer than Word 2. Unfortunately, Bat-El & Cohen (2012) do not provide the distribution of stress patterns for this kind of subset. Their constellation closest to ours is one where the blend's size is identical to the size of only one of the two polysyllabic base words. In this case Bat-El & Cohen (2012) predict two outcomes: if the blend has the same size as Word 2 (and Word 1 and Word 2 have different stress patterns), the blend is predicted to be stressed on the same syllable as Word 2. This prediction is largely correct (82 per cent of the pertinent cases in our data, N = 138). If, however, the blend has the same size as Word 1 (and Word 1 and Word 2 have different stress patterns), Bat-El & Cohen (2012) observe variable stress in their data. Either the blend is stressed on the same syllable as Word 1 (in 5 of their 11 pertinent cases), or on the same syllable as Word 2 (in 6 of their 11 cases). In our data we find 171 observations with this constellation, with 55 per cent of the blends preserving Word 1 stress. There is no significant difference between the two data sets ($\textit{chi-square} = 0.0087$, $df = 1$, $p = 0.93$). Note that this analysis relates the preservation of Word 1 stress crucially to the sizes of Word 1 and the blend. It is unclear whether this is a necessary or even useful idea, as we saw above that it is possible to make successful predictions without taking the size of the blend into account.

Another constellation in Bat-El & Cohen (2012) that has a bearing on the question of Word 1 vs Word 2 stress preservation and that is available for comparison with our data is the following. One of the base words is monosyllabic and the size of the blend is identical to the size of the longer word. For this constellation, Bat-El & Cohen (2012) find a near-categorical decision for a blend stress pattern that mirrors the stress pattern of the longer word. In our data, 78 per cent of the 262 pertinent observations behave in this way. Notably, if Word 1 is the longer one, the proportion of correct predictions is only 60 per cent (N = 116), while if Word 2 is the longer one, the proportion of correct predictions rises to 90 per cent. This means that in our data, the size effect for this particular constellation is weaker than in Bat-El & Cohen's (2012) data set, and, contrary to the prediction, Word 1 is not able to categorically preserve its stress in this

[4] As one reviewer correctly points out, Bat-El's blends are not all coordinative in nature, and this could in principle also lead to different results. However, this is rather unlikely since, to our knowledge, there is no study that has found systematic formal differences between coordinative and determinative blends.

constellation. Again it seems that introducing the size of the blend as a variable does not lead to satisfactory results.

With regard to the problem of monomorphemic, or what Bat-El & Cohen (2012) call 'default' stress, these authors see monomorphemic stress emerging under two particular circumstances. One is that with both formations the stressed syllables do not survive in the blend; the other is that one of the bases is monosyllabic and in the other base the stressed syllable does not survive. We will discuss each in turn.

As for the first constellation, in our data it is never the case that Word 1 and Word 2 both lose their main-stressed syllables. There are 95 cases (i.e. 13 per cent of a total of 709 blends with polysyllabic bases) in which Word 1 loses its stressed syllable, and 223 (31 per cent) in which Word 2 loses its stressed syllable. However, there is not a single case out of the 709 where both base words would lose their stressed syllable. This shows that speakers seem to systematically avoid such formations, and any other generalization over this class is probably misguided.

Let us now look at the other constellation for which Bat-El & Cohen (2012) predict default stress. This is the constellation where one of the bases is monosyllabic and in the other base the stressed syllable does not survive. In this case, default stress is claimed to be triggered 'regardless of the position of the monosyllabic word' (Bat-El & Cohen 2012: 208; unfortunately, the authors only give a few examples and no quantitative results). As we will see, our data do not at all support this claim.

We first discuss the subset in which Word 2 is monosyllabic and Word 1 is polysyllabic. This constellation is instantiated 177 times in our data set, with 105 items that have stress on the second or third syllable of Word 1, counting from the left. This subset of 105 is the one to test, because in AD blends the first syllable of Word 1 is preserved anyway, independent of stress. Strikingly, there is not a single case among these 105 items where Word 1 would lose the segmental material of its stressed syllable. Again the constellation for which Bat-El & Cohen (2012) evoke default stress is generally avoided by the speakers.

We finally consider the items where Word 1 is monosyllabic and Word 2 is polysyllabic, of which there are 170 observations in our data set. Of these we need to exclude 27 items where Word 2 is stressed on the final syllable, since this syllable must survive by definition, as we are dealing with AD blends. Of the remaining 143 items, 74 lose the stressed syllable of Word 2. Of these blends, 73 per cent show the same stress pattern as Word 2, which is in line with the idea developed above that the prosodic structure survives even if segmental material is lost. Only in the 27 per cent of the remaining cases do we find stress patterns that could be interpreted as 'default' stress assignment. However, quite often these patterns could also be accounted for by Word 1 stress preservation in the absence of the possibility of Word 2 stress preservation. For example, for *bárant* from *bar* + *restaurant*, an analysis is easily conceivable without reference to default stress: the dactylic stress pattern of Word 2 cannot be preserved, as the blend has only two syllables. In this case Word 1 stress is preserved.

To summarize our results on 'default' stress, we can say that the putative contexts for this are much more restricted than claimed by Bat-El & Cohen (2012). Two of the three contexts proposed by Bat-El & Cohen (2012) are actually never created, as the speakers prefer to preserve the prosodic patterns given by Word 2 or, if need be, Word 1. The only remaining context shows a very clear majority of Word 2 stress preservation.

Apart from the problems involved in their account of ‘default’ stress, there is also a more general problem with Bat-El & Cohen’s approach. The constellations they discuss are constellations in which crucially the size of the blend (in combination with certain sizes of the bases) is a determinant of stress assignment to the blend itself. Conceptually, this complicates matters considerably as it introduces two more correspondence relations into the model, and empirically it does not lead to convincing results.

Overall, Bat-El & Cohen’s (2012) approach seems to raise more questions than it answers, and it remains to be seen how the empirically observable patterns can be successfully implemented in an optimality-theoretic analysis.

5 Summary and discussion

This article has presented the first systematic quantitative investigation of the structure of experimentally elicited blends in English. It complements earlier quantitative work on the structure of existing, lexicalized blends (especially by Gries) in several ways. On a methodological level, it has shown that experimentally elicited blends constitute a valid resource for research on blend structure, which crucially allows for the investigation of structural variation both within and across types for controlled structural configurations. Importantly, we find that our experimental blends exhibit structural properties which have been documented for existing blends in the literature (e.g. strong preference for AD structure, generalizations on length, and various others). With respect to the phenomenon under scrutiny, the present study has extended earlier analyses in that it has looked at a variety of ways in which prosodic factors influence blend structure in English. Earlier work concentrated on only some of such factors. For example, Gries (2004a, 2004b, 2006, 2012) looks at length as the only prosodic factor considered, and Bat-El & Cohen (2012) exclusively focus on blend stress. The results of the present analysis suggest that prosody has a much more far-reaching impact on blend structure than hitherto noted for English. This not only lends support to the classification of blending as an instance of so-called prosodic morphology in some of the earlier literature (as, e.g., in Plag 2003), but also has important implications for general approaches to blend structure that have been proposed in the theoretical literature. These will be discussed in the remainder of this section.

We systematically investigated three aspects of blend structure: the length of the blend, the location of the switchpoint, and the stress of the blend. Our analysis suggests that these aspects are crucially determined by the prosodic properties of the two base words of the blend. Of all prosodic categories studied, stress in the base word emerged as the most important factor. This is especially true for stress in Word 2, which was shown to have a crucial influence not only on the stress pattern of the blend, but also on the location of the switchpoint. Thus, we see a clear tendency for switchpoints to be located in the main-stressed syllable of Word 2. A plausible interpretation of this fact is that blends are optimized with regard to the recoverability of the two base words by preserving exactly those parts of the two base words that provide important cues to word recognition (see Gries 2006 for a similar argument). In the case of Word 1, this is word-initial material; in the case of Word 2, this is material that has to do with the position of main stress. With respect to the latter generalization, our analysis showed that stress preservation effects may in fact be more complex than hitherto noted. Thus, whereas it is true that the switchpoint tends to be in the stressed syllable of Word 2, this is not without exceptions. It is these exceptions that provide us with an important insight about the nature of stress preservation effects, which is that stress preservation may happen independently of the preservation of segmental material of the stressed syllable (see Piñeros 2000, who develops a similar argument, but on different grounds). Two illustrative cases are the blends *jóurnarator* (from *jóurnalist* and *móderator*) and *préstinant* (from *prestígious* and *dóminant*). In terms of segmental content, the switchpoint in these blends is not the stressed syllable of Word 2. The blends preserve less material from Word 2 than the string from the stressed syllable to the right word edge. However, in terms of the stress pattern, the stressed syllable of Word 2 is represented in the blend in so far as the blend preserves the location of the main-stressed syllable of Word 2, i.e. independently of its segmental content. In the case of *jóurnarator* this stressed syllable corresponds to the stressed syllable of Word 1 (*jour* in *jóurnalist*), in the case of *préstinant* it does not (*pres* is unstressed in *pres.tí.gious*).

In contrast to stress, and contrary to earlier claims, syllabic constituency was shown to be of minor importance for switchpoint location. Length of the source words seems to have an influence on blend structure in that blend length most often conforms to the length of the longer source word. At the same time, there seems to be a preference for blends to have no more than three syllables.

The important role of stress in the base word for the structure of blends raises the question of how our findings relate to other, prosody-based accounts of blend structure that can be found in the theoretical literature. One often-cited account is the proposal put forward in Piñeros (2004) for a blend pattern found in Colombian Spanish. According to this analysis, blending involves the integration of one base word into the prosodic structure of the other word. A representative example is the blend [ˌʧib.ʧom.ˈbja.no] (‘a Colombian of Chibchan descent’) from the base words [ˈʧib.ʧa] (‘Chibchan’) and [ˌko.lom.ˈbja.no] (‘Colombian’), where the length and prosodic structure of the blend fully corresponds to the length and the prosodic structure of Word 2. Our findings for English blends suggest that it is indeed true that, as shown above, preservation of the prosodic structure of Word 2 is an important factor. However, they are not fully compatible with the idea that the prosodic structure of Word 2 acts as a template into which the surviving part of Word 1 is integrated. Instead, a distinction must be made between pretonic and post-tonic blend structure. As we have seen, stress in

Word 2 acts as a major determinant of stress in the blend and of the location of the switchpoint. Pretonically, however, prosodic structure and length of Word 2 does not determine blend structure to the same extent. Instead, we find variation. A representative example of such variation is the group of blends that our experimental subjects coined for the base words *pílot* and *ófficer*. Of the 17 pertinent tokens, only 7 tokens are compatible with Piñeros' account for Spanish. These are *pílicer*, *píficer* and *pófficer*. By contrast, eight subjects coined *pilófficer*, which faithfully preserves the prosodic structure of Word 2 only from the stressed syllable onwards. This type of variation is not compatible with approaches that assume that blend shape is determined by a prosodic template that demands blends to conform to the Word 2 skeleton in a categorical manner (as proposed by Piñeros 2000 for Spanish; see also Trommer & Zimmermann 2012). Instead, it suggests that the preservation of Word 2 prosodic structure is a requirement that may be violated in systematic ways, thus allowing for variation to occur.

The influence of prosodic factors on blend structure that our analysis has brought to light opens up new avenues for further research. For example, it remains to be seen how the apparent interaction of stress-related constraints with other restrictions can be formally modeled, and how the patterns of variability and non-variability observable in the data may emerge in such a model of blend formation.

Authors' addresses:
Institut für Anglistik und Amerikanistik
English Language and Linguistics
Heinrich-Heine-Universität Düsseldorf
40204 Düsseldorf
Germany
sabine.arndt-lappe@uni-duesseldorf.de
ingo.plag@uni-duesseldorf.de

References

Alber, Birgit. 2001. Maximising first positions. In Caroline Féry & Ruben van der Vijver (eds.), *Proceedings of HILP 5*, 1–19. Potsdam: Universitätsverlag Potsdam.

Alber, Birgit & Sabine Arndt-Lappe. 2012. Templatic and subtractive truncation. In Jochen Trommer (ed.), *The phonology and morphology of exponence – the state of the art*, 289–325. Oxford: Oxford University Press.

Baayen, R. Harald. 1992. Quantitative aspects of morphological productivity. *Yearbook of Morphology 1991*, 109–49.

Baayen, R. Harald. 1993. On frequency, transparency and productivity. *Yearbook of Morphology 1992*, 181–208.

Bat-El, Outi. 1996. Selecting the best of the worse: The grammar of Hebrew blends. *Phonology* 13, 283–328.

Bat-El, Outi. 2006. Blend. In Keith Brown (ed.), *Encyclopedia of language & linguistics*, 2nd edn, 66–70. Amsterdam: Elsevier.

Bat-El, Outi & Evan-Gary Cohen. 2012. Stress in English blends: A constraint-based approach. In Vincent Renner, François Maniez & Pierre J. L. Arnaud (eds.), *Cross-disciplinary perspectives on lexical blending*, 193–212. Berlin: Mouton de Gruyter.
Bauer, Laurie. 1983. *English word-formation*. Cambridge: Cambridge University Press.
Bauer, Laurie. 2012. Blends: Core and periphery. In Vincent Renner, François Maniez & Pierre J. L. Arnaud (eds.), *Cross-disciplinary perspectives on lexical blending*, 11–22. Berlin: Mouton de Gruyter.
Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
Ben Abdallah, Manel. 2008. The structural predictability of English blends: A case study of novel blend formation in Northern Ireland English. Master's thesis, University of Siegen.
Benua, Laura. 2000. *Phonological relationships between words*. New York: Garland.
Cannon, Garland. 1986. Blends in English word formation. *Linguistics* 24, 725–53.
Costa, Nicoletta. 2008. *Das Phänomen der Wortkreuzung im Deutschen*. Verona: Università degli Studi die Verona Tesi di laurea specialistica.
Fischer, Roswitha. 1998. *Lexical change in Present-Day English*. Tübingen: Narr.
Gries, Stefan T. 2004a. Isn't that fantabulous? How similarity motivates intentional morphological blends in English. In Michael Achard & Suzanne Kemmer (eds.), *Language, culture, and mind*, 415–28. Stanford: CSLI Publications.
Gries, Stefan T. 2004b. Shouldn't it be *breakfunch*? A quantitative analysis of the structure of blends. *Linguistics* 3(42), 639–67.
Gries, Stefan T. 2006. Cognitive determinants of subtractive word-formation processes: A corpus-based perspective. *Cognitive Linguistics* 17(4) in press.
Gries, Stefan T. 2012. Quantitative corpus data on blend formation: psycho- and cognitive-linguistic perspectives. In Vincent Renner, François Maniez & Pierre J. L. Arnaud (eds.), *Cross-disciplinary perspectives on lexical blending*, 145–67. Berlin: Mouton de Gruyter.
Kubozono, Haruo. 1990. Phonological constraints on blending in English as a case of phonology-morphology interface. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1990*, 1–20. Dordrecht: Foris.
Lappe, Sabine. 2007. *English prosodic morphology*. Dordrecht: Springer.
Marchand, Hans. 1969. *Categories and types of Present-Day English word-formation*. Munich: C. H. Beck.
McCarthy, John J. & Alan Prince. 1993. *Prosodic morphology I: Constraint interaction and satisfaction* (Technical Reports of the Rutgers Center for Cognitive Science 3). New Brunswick, NJ: Rutgers University.
McCarthy, John & Alan Prince. 1998. Prosodic morphology. In Andrew Spencer & Arnold Zwicky (eds.), *The handbook of morphology*, 283–305. Oxford: Blackwell.
Piñeros, Carlos E. 2000. Prosodic and segmental unmarkedness in Spanish truncation. *Linguistics* 38, 63–98.
Piñeros, Carlos E. 2004. The creation of portmanteaus in the extragrammatical morphology of Spanish. *Probus* 16(2), 203–40.
Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin and New York: Mouton de Gruyter.
Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.
Trommer, Jochen & Eva Zimmermann. 2012. Portmanteaus as generalized templates. In Vincent Renner, François Maniez & Pierre J. L. Arnaud (eds.), *Cross-disciplinary perspectives on lexical blending*, 233–58. Berlin: Mouton de Gruyter.

Appendix 1: Stimuli

1. This fruit has a ______________________ (sour+sweet) flavour.
2. Dad bought a ______________________ (scanner+printer) machine.
3. He has a ______________________ (Turkish+Polish) origin.
4. It is the first time since the end of the Cold War that a ______________________ (fighter+bomber) is being regularly scrambled from UK air bases to monitor aircraft in national airspace.
5. Being a ______________________ (daughter+mother), she has to be mature.
6. Joe Strupp is a senior ______________________ (publisher+editor).
7. He has been working as a ______________________ (journalist+moderator) since 2000.
8. Researches which are concerned with the ______________________ (modernist+feminist) theory are growing.
9. I saw a ______________________ (terrible+horrible) accident in the news yesterday.
10. US scientists discover a new plant which resembles a ______________________ (tomato+potato).
11. Kenneth Laurence is an American ______________________ (singer+actor).
12. They aim to ______________________ (modify+simplify) the managerial functions of the company.
13. ______________________ (employer+director) is a representative of his respective constituent.
14. The wall is painted in a ______________________ (blue+green) colour.
15. My mother is responsible for the ______________________ (consultation+ interpretation) of private and business clients.
16. The album is full of ______________________ (black+white) photos.
17. Leontief, the innovator of the ______________________ (input+output) model, won a Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel for his development of this model.
18. Keiko lent her ______________________ (Japanese+Chinese) dictionary to her friend.
19. Titanic was a ______________________ (gigantic+enormous) boat.
20. Mum uses a ______________________ (spoon+fork) to mix the salad.
21. People talk about the expansion of ______________________ (sex+alcoholism) throughout the world.
22. She enjoys eating her ______________________ (bread+butter) sandwich.
23. The teacher asks the pupils to determine the ______________________ (function+form) of the words.
24. Brian has a part-time job as a ______________________ (cook+waiter) in a small restaurant.
25. This is a ______________________ (seminar+lecture) for postgraduate students.
26. ______________________ (Austria+Hungary) is a dual-monarchic union state in central Europe.
27. They decided to go to a ______________________ (bar+restaurant).
28. The company specializes in producing. ______________________ (plastic+wood) kitchen tools.
29. The ______________________ (export+import) Bank supports the financing of US goods and services.
30. She wanted a new ______________________ (sofa+bed) for her birthday.
31. There is not enough space in the room so she bought a ______________________ (table+desk).
32. For the English course, students have to write an ______________________ (essay+text).

33. The play is based on ______________________ (sarcasm+catastrophe).
34. Mrs James is a University ______________________ (lecturer+tutor).
35. I have a ______________________ (photocopier+printer) at home.
36. He likes the ______________________ (hunter+gatherer) way of life.
37. Small families tend to live in an ______________________ (apartment+house).
38. Walter Lippmann is a famous ______________________ (writer+photographer).
39. They offer a ______________________ (German+Portuguese) translation service.
40. She prefers a ______________________ (coffee+milk) drink in the morning.
41. A ______________________ (dictionary+thesaurus) helps students improve their vocabulary.
42. The secret of an ______________________ (admirer+lover) is unbreakable.
43. He applies for a job as an ______________________ (executor+manager).
44. To cook the pasta they use a large ______________________ (pan+pot).
45. The new Winter–Autumn collection includes a new type of a ______________________ (jacket+coat) combination.
46. Miss Bill has her own ______________________ (studio+apartment).
47. The boundaries of ______________________ (south+east) England include Berkshire, Buckinghamshire, East Sussex, Hampshire, Isle of Wight, Kent, Oxfordshire, Surrey and West Sussex.
48. Students who live on campus have a ______________________ (shower+bath).
49. Via Mercanti leads off from the ______________________ (north+ west) corner of the square.
50. David Lewin is a ______________________ (pilot+officer).
51. They still retain their stronghold in ______________________ (south+west) Scotland and northern England today.
52. Got a ______________________ (girl+friend)?
53. Asked if he'd been her ______________________ (boy+friend) for long, he said 'no'.
54. By the end of January 1946 only two ______________________ (British+Indian) battalions remained.
55. The ______________________ (secretary+general) has been asked to produce his report on a range of subjects by 1 July.
56. The strongest security was provided on the ______________________ (Slovak+Ukrainian) border.
57. ______________________ (Lords+chancellors) have senior and important functions in the government of the United Kingdom.
58. The ______________________ (Lord+chief) Lord Phillips of Worth Matravers has challenged magistrates in Dorset to support efforts to speed up the administration of criminal justice and to assist the process by finding alternatives to custody as a means of punishing offenders.
59. The UAB ______________________ (spouse+partner) Relocation Program offers relocating spouses and partners of newly recruited UAB employees assistance with the job search process.
60. The most ______________________ (prestigious+dominant) site in Belfast is thought to be a prismatic site on Ann Street, which normally works out at £350 for two weeks.

Appendix 2

Table A1. *Switchpoint by compatibility of onsets (N = 100)*

	Switchpoint at constituent boundary	Switchpoint within onset
Compatible onsets	28	12
Non-compatible onsets	60	0

Table A2. *Switchpoint by Word 2 stress (N = 1263)*

	Word 2 stress, number of syllables from right				
Switchpoint Word 2, number of syllables from right	1	2	3	4	5
1	406	108	18	1	0
2	24	297	134	7	1
3	3	52	165	8	10
4	0	5	3	3	11
5	0	1	0	1	3
6	0	0	0	1	0
7	0	0	0	1	0

Table A3. *Blend stress by Word 2 stress*

	Stressed syllable in W2, from right				
Stressed syllable in blend, from right	1	2	3	4	5
1	34	1	1	0	0
2	3	315	38	12	0
3	3	45	227	4	2
4	0	3	10	6	0
5	0	0	1	0	0

Table A4. *Blend stress by Word 1 stress*

	Stressed syllable in W1, from left		
Stressed syllable in blend, from left	1	2	3
1	321	26	3
2	139	83	5
3	72	25	16
4	3	1	10
5	0	0	1