OTHER PAPER

Written production of German compounds

Effects of lexical frequency and semantic transparency

Said Sahel[1], Guido Nottbusch[1], Angela Grimm[2] and Rüdiger Weingarten[1]
[1]University of Bielefeld, Germany / [2]University of Frankfurt/Main, Germany

In this study, we present an experiment in which we examined the time course of typing German compounds. The compounds varied according to three criteria: (1) whole word frequency (high vs. low), (2) head frequency (high vs. low) and (3) semantic transparency (transparent vs. opaque). In this experiment, we recorded the interkey intervals (IKIs) and concentrated on the IKI measurements found at the boundary of the two immediate constituents in compounds. We refer to this boundary type as an SM-boundary because (S)yllable and (M)orpheme boundaries coincide at this word position. As we found effects of lexical frequency for SM-IKIs in a series of previous studies, we argue that possible differences in SM-IKIs found for compounds of different frequency classes and of different degrees of semantic transparency can give an insight into the processes involved in the written production of German compounds: whole word procedures and/or compositional procedures. Our findings show that SM-IKIs are affected by compound frequency, head frequency and semantic transparency. We therefore argue that both whole word procedures and compositional procedures are involved in the written production of German compounds. These findings are in line with those versions of dual-route models which postulate that the two routes run in parallel and interact.

1. Introduction

The issue of how morphologically complex words are mentally represented and accessed has been a widely discussed topic in psycholinguistic literature, the central question being: Are morphologically complex words represented and accessed as whole words or in terms of their constituents? There are differing views on how polymorphemic words are represented and accessed. They range from

Written Language & Literacy 11:2 (2008), **211–227**. DOI 10.1075/wll.11.2.06sah
ISSN 1387–6732 / E-ISSN 1570–6001

models postulating morphological (de)composition for all polymorphemic words (e.g. Taft & Forster, 1975) to full-listing models (e.g. Butterworth, 1983). In this context, the distinction between compositional and non-compositional models was established. According to compositional models, the production and recognition of polymorphemic words require morphological composition and decomposition of their constituents, respectively. Non-compositional models, however, claim that polymorphemic words are represented and accessed as full forms, just as monomorphemic words. None of these extreme theoretical positions can provide adequate explanations for the existing empirical findings. These may only be accounted for by the consideration of models suggesting whole word access as well as (de)composed access and representations for polymorphemic words. Therefore, these dual-route models have become widely accepted over the last fifteen years (see the Augmented Addressed Morphology model of Caramazza and colleagues: e.g. Caramazza, Laudanna & Romani, 1988; Chialant & Caramazza, 1995; the Morphological Race Model of Frauenfelder & Schreuder, 1991; the meta-model of Schreuder & Baayen, 1995; for a review see McQueen & Cutler, 1998). Dual-route models claim that polymorphemic words can be accessed both by the activation of whole-word units and by the activation of their constituent morphemes.

As mentioned above, dual-route models posit that the processing of polymorphemic words involves the whole word route as well as the constituent route. These models also suggest that the manner in which a polymorphemic word is processed is crucially affected by its lexical properties such as frequency and semantic transparency. As these models have been constructed almost exclusively on the basis of data taken from studies of language comprehension, the assumptions they make about the processing of polymorphemic words have not been satisfactorily verified for the field of language production. Therefore, in order to clarify the mechanisms involved in the production of polymorphemic words, research needs to be carried out on language production. To our knowledge, only two studies have been carried out concerning the production of compounds by normal adults (Dohmes, Zwitserlood & Bölte, 2004; Bien, Levelt & Baayen, 2005). Dohmes, Zwitserlood & Bölte (2004) conducted a series of picture naming tasks to assess the impact of semantic transparency on the naming latencies of German compounds. Contrary to their expectations, the authors found that semantically transparent distractors did not accelerate picture naming latencies compared to semantically opaque ones. The authors found that transparent distractors (e.g. ‹Wildente› ‘wild duck’), used as written primes, facilitated picture naming (e.g. picture of a duck) to the same degree as opaque ones (e.g. ‹Zeitungsente› ‘false report’, literally ‘newspaper duck’) and came to the conclusion that semantic transparency has no critical impact on the production of German compounds in spoken language. As their findings are inconsistent with the available evidence from language comprehension studies which did find effects of semantic transparency in the recognition of compounds (Sandra, 1990; Zwitserlood, 1994; Libben, 1998; Libben, Gibson, Bom Yoon & Sandra, 2003), Dohmes, Zwitserlood & Bölte (2004: 211) assume that semantic transparency "might play a different role in language production than in comprehension".

Bien et al. (2005) investigated the impact of constituent frequency and compound frequency on the production latencies of Dutch compounds in a series of spoken naming experiments. They found that the frequencies of both initial and second constituents affect the production latency of compounds. Compounds with a high initial constituent frequency elicit shorter latencies than those with a low initial constituent frequency and compounds with a high second constituent frequency elicit shorter latencies than those with a low second constituent frequency. As regards the impact of the compound frequency, the authors found no effect on production latencies. The latencies of high-frequency compounds were no faster than the latencies of low-frequency ones. In a further analysis, the authors found indications that compound frequency accelerates the naming latencies in the lower range and slows them down in the higher range of compound frequencies. Based on the clear effects of the constituent frequencies on the naming latencies of compounds, the authors argue that their findings support decompositional models of compound production.

The processing of compounds in language production is best examined in aphasics (e.g. Delazer & Semenza, 1998; Blanken, 2000; Badecker, 2001). All these studies focus on the issue of whether aphasics make use of compositional procedures in the production of compounds. The performance and error patterns in the naming tasks conducted in these studies suggest that the production of compounds involves lexical composition. Blanken (2000) reported on a group study with twenty German-speaking aphasics who exhibited difficulties in producing compounds. The naming errors often resulted from the substitution of one constituent or from the simplification of compound targets. According to the author, the observed error patterns suggest that, in keeping with dual route models, both the whole word route and the constituent route are involved in the production of compounds. Moreover, Blanken (2000) found effects of lexical frequency and semantic transparency on compound production: the naming accuracy rate for compounds made up of two highly frequent constituents was higher than for those with two low frequency constituents and higher for transparent than for opaque compound targets. Blanken (2000) also discovered that certain errors were rarely made with opaque but often with transparent compounds. This indicates that whole word procedures were to a greater extent involved in the production of opaque compounds than in the production of transparent compounds.

Further neurolinguistic evidence supporting the involvement of compositional mechanisms in compound production comes from a single case study in which Badecker (2001) reported on the performance of an English-speaking aphasic (CSS) in two naming tasks with monomorphemic and compound targets. The author found a performance asymmetry for monomorphemic and compound targets concerning both the naming accuracy and the error patterns. On the one hand, CSS gave more correct responses to monomorphemic targets than to compound targets whereas, on the other hand, certain substitution and misordering errors were observed only for compound targets. Badecker (2001) claims that this asymmetry indicates that the lexical system takes into account the morphological structure of the words in question. He also argues that these findings can only be explained by the assumption that CSS's compound production includes compositional procedures. In contrast to Blanken (2000), however, no lexical frequency effects were observed: the accuracy rate of CSS's compound naming was neither affected by whole word frequency nor by the frequency of the constituent morphemes.

In a further single case study, Delazer & Semenza (1998) reported on an Italian-speaking aphasic (MB) whose ability to name objects inducing compounds was selectively disturbed. They found that the compound structure was largely preserved in MB's errors with the most frequent error type being substitution errors where one constituent was replaced while the second was preserved. At the same time, the preserved constituent kept the position within the compound. However, no substitutions of compounds by monomorphemic words were observed. The authors infer that MB's compound production involves the activation of the two morpheme constituents and this is taken as evidence supporting compositional procedures. Similar to Badecker (2001), no lexical frequency effects were observed.

In the present study, we examine the time course of German compound production in written language and we address two main issues: 1) does the production of German compounds require compositional and/or whole word procedures? and 2) do factors such as lexical frequency and semantic transparency determine the processing manner? These issues will be assessed using a single word typing paradigm that has already been employed in a number of recent studies which explore the impact of linguistic units, such as syllables, morphemes and graphemes, in the time course of typing (Weingarten, Nottbusch & Will, 2004; Sahel, Nottbusch, Blanken & Weingarten, 2005, Nottbusch, Grimm, Will & Weingarten, 2005; Will, Nottbusch & Weingarten, 2006; Nottbusch, Weingarten & Sahel, 2006). All of these studies have shown that the time course of typing is crucially influenced by the linguistic structure of words and that the interkey intervals (henceforth IKIs) for identical bigrams, i.e. the delays between two keystrokes (e.g. <l-s>), vary significantly depending on the type of linguistic boundary: IKIs at a syllable boundary (henceforth S-IKIs) in words like ‹Fel-sen› ‘rock’ are longer than within-syllable IKIs (henceforth L-IKIs) in words like ‹fal-sch› 'wrong' and IKIs where a morpheme and a syllable boundary coincide (henceforth SM-IKIs) in words like ‹Roll-schuh› 'roller-skate' are longer than S-IKIs in words like ‹Fel-sen› 'rock'. However, no effects were found at 'pure' morpheme boundaries (henceforth M-IKIs), i.e. within-word positions where a morpheme boundary and a syllable boundary do not coincide in words like ‹Kind-er› 'children'. Here, the studies revealed no significant differences between M-IKIs and L-IKIs. A further central finding of these studies concerns the impact of the lexical frequency on the time course of typing. Although highly frequent words were typed faster in general, frequency effects were found only for SM-IKIs: SM-IKIs of highly frequent words were shorter than the SM-IKIs of low-frequency words (Weingarten, Nottbusch & Will, 2004: 539). In all other within-word positions, including syllable boundaries, no effects of lexical frequency were revealed. According to the common view that word frequency effects indicate lexical access and that "lexical processes operate over units at least the size of a morpheme" (Rapp, Epstein & Tainturier, 2002: 1), we assume that lexical processes become operative at SM-boundaries. In general, these findings are in line with results from previous studies that concern the effects of the morphological structure of words on the time course of writing (handwriting and typing), indicating that morphemes can be seen as processing units in writing (e.g. Pynte, Courrieu & Fenck, 1991; Orliaguet & Boë, 1993).

The present study measures the delays found in typing compounds and specifically concentrates on SM-type boundaries. In our case, a boundary of the SM-type is identical with the right-hand boundary of the first and the left-hand boundary of the second immediate constituent, i.e. the head of a compound. We therefore measure SM-IKIs, i.e. the intervals between the last letter/key of the first immediate constituent and the first letter/key of the second immediate constituent of compounds. In so far as our previous studies have shown that SM-IKIs are strongly affected by lexical frequency which indicates lexical access, we assess the above issues by analyzing IKIs at SM-boundaries. These are scrutinized in order to determine whether delays at the SM-boundaries of compounds are affected by compound frequency and/or by head frequency. Compound frequency effects would provide evidence for a holistic pathway, whereas head frequency effects could be viewed as supporting a compositional pathway in the written production of compounds. Secondly, we examine the influence of semantic transparency on SM-IKIs. Effects that arise from different SM-IKIs for semantically transparent and semantically opaque compounds would provide evidence for different processing procedures involved in the written production of compounds.

2. The study

In order to test the influence that lexical frequency and transparency might have on the time course of the typing of compounds, we conducted an experiment in which 45 German participants were asked to type compounds on a computer keyboard. The compound targets had different frequency levels and different degrees of semantic transparency.

2.1 Method

Participants: 45 students (34 female and 11 male) from the University of Osnabrück participated in the experiment. All were native speakers of German. They were between 22 and 38 years of age (Mean: 25.9, SD: 3.6). Almost all participants were right handed (42), only three were left handed. All participants were able to type fluently without hesitation, although no strict criterion was applied to the typing speed.

Stimuli: The stimulus material consisted of 168 German compounds with word length ranging from 7 to 27 characters (Mean: 10.7, SD: 2.6): 131 noun-noun, 18 verb-noun, 16 adjective-noun, two preposition-noun and one particle-noun compounds. Apart from 4 items, the first immediate constituent of all compounds was monomorphemic. In 128 cases, the head was a monomorphemic noun, whereas in 40 cases it consisted of two or three morphemes. All compounds varied according to three criteria: 1) compound frequency (high vs. low), 2) head frequency (high vs. low) and 3) semantic transparency, i.e. whether or not the meaning of the compound is derivable from the meaning of its two immediate constituents (transparent vs. opaque). For the purpose of this study, only SM-IKIs, i.e. the intervals between the last letter/key of the first immediate constituent and the first letter/key of the head of compounds, were of interest. Therefore compound targets were arranged in a way that groups of 6 to 12 words a) shared the same bigram at the SM-boundary and b) featured different levels of the variables in question, if possible.

The level of semantic transparency was determined by five linguistically skilled academic researchers. This led to a list of 315 German compounds containing one of 23 different bigrams at the SM-boundary, i.e. at the boundary between the two immediate constituent morphemes of the compound. The subjects were asked to judge the transparency level of all compounds on a five-point scale with 168 experimental items being selected from the extreme ends of this rating. The numbers of cases in each class was fairly equal (70 opaque vs. 98 transparent).

Frequency measurements in terms of natural log [log (freq +1)] were taken from the CELEX database (Baayen, Piepenbrock & van Rijn, 1993). Items not contained in CELEX or having an absolute frequency of zero were assessed with a log value of zero. The log of compound frequencies ranged from 0 to 6.55 with a mean of 1.33 (SD: 1.72). Head frequencies, on the other hand, had a range of 7.87 and a mean value of 3.32 (SD: 2.29). The varying frequency distribution poses a problem judging the borderline between high and low frequency values. There are two possible solutions: (1) The first solution is to set a different borderline for both types of frequencies in order to have equal numbers of cases in all groups (compound frequency: high and low, head frequency: high and low). (2) The second solution is to accept the loss of equally distributed numbers of cases in favour of a single borderline for both types of frequency. We chose the second option. The borderline between the high and low frequency level was set at 2.75 (corresponding to an absolute frequency of 15.64 per 5.4 million words of the Mannheim corpus of written texts contained in CELEX, i.e. items with an absolute frequency from zero to 15 were rated as low frequency items and items with an absolute frequency from 16 to the highest value were rated as high frequency items). By doing this, 79.2% of the compound frequencies (133 low frequency compounds vs. 35 high frequency compounds) and 46.4% of the head frequencies (78 low frequency heads vs. 90 high frequency heads) were rated as low frequency items, i.e. the sum of the absolute differences to 50% (here: ABS(50–79.2) + ABS(50–46.4) = 32.8) was lowest in setting the borderline to 2.75. The combination of the three variables led to the distribution as shown in Table 1, which also reflects some language constraints: the number of high-frequency compounds is quite limited, especially in cases of semantic transparency and even more so if there is a low frequency head. However, low-frequency compounds, especially transparent ones, are numerous.

The 168 compound targets used in this experiment contained 23 different bigrams at the SM-boundary, i.e. at the boundary between the two immediate constituents of a compound. Our basic assumption is that a fluent typist will, in all likelihood, type identical bigrams in the same motor pattern, i.e. each key with the same finger, independent of the learned typing system (touch-typing vs. two-finger-typing). 18 bigrams, namely ‹df›, ‹ds›, ‹gg›, ‹hb›, ‹hh›, ‹lk›, ‹mb›, ‹nw›, ‹rf›,

Table 1. Distribution of compounds by whole word frequency, head frequency and semantic transparency.

				semantic transparency		Σ	
				transparent	**opaque**		
compound frequency	high	head frequency	high	8	10	18	35
			low	5	12	17	
	low	head frequency	high	44	28	72	133
			low	41	20	61	
			Σ	98	70	168	

‹rh›, ‹rk›, ‹rm›, ‹rs›, ‹sb›, ‹te›, ‹tf›, ‹tl›, tn›, occurred in 6 words each, and 5 bigrams, namely ‹ns›, ‹ta›, ‹tg›, ‹tk›, ‹ts›, occurred in 12 words. Each group of items containing the same bigram consisted of at least one item of both compound frequency classes, one item of both head frequency classes and one item of both transparency classes. However, due to language constraints (see above) it was not possible to have all possible combinations of variables for all bigrams.

Apparatus: The experiment was conducted using a standard PC with a 19”-CRT display. Stimulus display and keystroke measurements were controlled by ExpKit, a program designed to record keyboard data with maximal sampling rate in psycholinguistic experiments written by Boris Gutbrod. By using this equipment, we were able to achieve an accuracy of approximately 8 ms.

Procedure: The participants were given detailed written instructions before testing. Each trial started with the display of a blank grey screen for 1,000 ms, followed by the presentation of an asterisk in the upper half of the screen for 300 ms at a position where the first letter of the stimulus was to appear. Simultaneously, a short beep was sounded. Following a blank screen of 200 ms, the stimulus word was presented. Participants were instructed to carefully read the stimulus and then type the word on the keyboard as fast as possible without errors. On typing the first letter of the target word, the stimulus disappeared from the screen. The next trial was initiated by the participant and started 1,500 ms after pressing the return key. Eight training items (not contained in the main list) were presented before the main test. Following the pre-test, and if necessary repeated instructions, the main test was conducted. All stimuli were randomly presented.

2.2 Results

Mistyped words (13.2%) and values exceeding 2.5 standard deviations of the mean IKI of the participant/item (4.2%) were discarded. The remaining 6239 single measurements (compound frequency: 1311 high vs. 4928 low; head frequency 3396 high vs. 2843 low; semantic transparency: 3588 transparent vs. 2651 opaque) were considered in the analysis. The average writing speed of the participants for the complete words was 46.0 five letter words/min (SD: 8.4) or 271.5 ms (SD: 74.8) mean Interkey Interval.

In order to compensate for the unequal numbers of cases and to incorporate the influence of different bigrams and individuals, a mixed model ANOVA analysis with the fixed factors COMPOUND FREQUENCY (high vs. low), HEAD FREQUENCY (high vs. low) and SEMANTIC TRANSPARENCY (transparent vs. opaque) as well as the random factors BIGRAM (23) and PARTICIPANTS (45) was used. In doing this, differences in bigrams and level of typing ability are dealt with by subtracting the bigram’s and the participant’s mean from each observation, i.e. the scores were

Table 2. Mixed model ANOVA results for IKI at SM-boundaries of German compounds (full model)

source	DF1	DF2	F-value	significance
COMPOUND FREQUENCY	1	6173	11.48	< .001
HEAD FREQUENCY	1	6185	4.56	< .05
TRANSPARENCY	1	6181	4.46	< .05
COMPOUND FREQUENCY * HEAD FREQUENCY	1	6183	< 1	> .5
COMPOUND FREQUENCY * TRANSPARENCY	1	6179	4.37	< .05
HEAD FREQUENCY * TRANSPARENCY	1	6179	< 1	> .5
COMPOUND FREQUENCY * HEAD FREQUENCY * TRANSPARENCY	1	6185	1.40	> .1

transformed to distances from the bigram and participant mean. The results for the fixed effects in the full model can be observed in Table 2.

Table 2 shows a highly significant *p* value for the main factor COMPOUND FREQUENCY and a significant effect for the main factors HEAD FREQUENCY and TRANSPARENCY. No interactions display effects, except for COMPOUND FREQUENCY * TRANSPARENCY. The estimated variance for the random factor BIGRAM was 5053.5

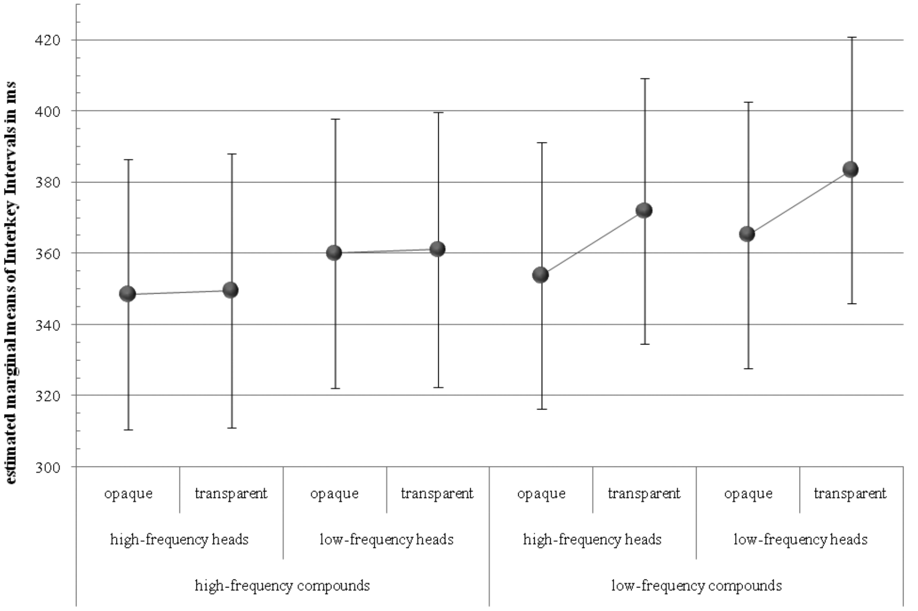

Figure 1. Estimated marginal means and error bars (95% confidence interval) for inter-key intervals at the boundaries of the two immediate constituents of German compounds. Note that the values are predicted from an optimal model (backwards selection; see text), not by observed means.

(SE: 1564), 5125.3 (SE: 1112) for the random factor PARTICIPANTS and 3402.9 (SE: 263) for the interaction between the two. In order to find the optimal model (i.e. identify the factors with the most significant effects in the full model), a step-by-step reverse selection was employed which resulted in the removal of the most non-significant variable. As can be observed in Figure 1, transparency only affects low-frequency compounds.

3. Discussion

The main aim of this study was to assess the influence of lexical frequency and semantic transparency on the production of German compounds by analyzing the time course of typing in skilled writers. We hypothesized that different SM-IKIs, i.e. the time elapsed during the transition from the last letter/key of the first immediate constituent to the first letter/key of the second, found for compounds of different frequency classes and of different transparency/opacity degrees would give an insight into the processes involved in the written production of German compounds: whole word procedures and/or procedures including constituent morphemes. Our study is, to our knowledge, the first of its kind on written language production dealing with the impact of lexical frequency and semantic transparency in compound production.

Our results indicate that both lexical frequency and semantic transparency affect the time course of the typing in German compounds. We found that SM-IKIs were affected by compound frequency, head frequency and semantic transparency with the SM-IKIs of high-frequency compounds being shorter than those of low-frequency ones, irrespective of their head frequency (i.e. whether they contain a high- or low-frequency head). Head frequency also plays a crucial role in the time course of compound production: a high-frequency head decreases the duration of SM-IKIs of high-frequency as well as low-frequency compounds. Within each of these two frequency classes of compounds, SM-IKIs in compounds with a high-frequency head were typed faster than those of compounds with a low-frequency head. The constant effects of compound and head frequency resulted in the following hierarchy of SM-IKIs, which begins with the fastest SM-IKIs:

– SM-IKIs of high-frequency compounds with a high-frequency head,
– SM-IKIs of high-frequency compounds with a low-frequency head,
– SM-IKIs of low-frequency compounds with a high-frequency head,
– SM-IKIs of low-frequency compounds with a low-frequency head.

These results suggest that compound frequency and head frequency have cumulative effects on SM-IKIs (no interaction). We found that, when a high compound frequency was combined with a high head frequency, the fastest SM-IKIs were produced whereas a low compound frequency combined with a low head frequency resulted in the slowest SM-IKIs. The frequency effects of both compounds and heads found at SM-boundaries, i.e. the transition from the first to the second immediate constituent of a compound, can only be accounted for by assuming that the lexical units involved are active at this stage of compound production. According to the assumption that frequency effects indicate lexical access (e.g. Jescheniak & Levelt, 1994), we can assume that lexical processes become operative at boundaries of the SM-type. Furthermore, it is plausible to conclude that the lexical unit corresponding to the compound is accessed when compound typing is initiated and remains active until the production of the second immediate constituent is, at least, initiated. It is only by assuming this that we can account for the effects of compound frequency found at the SM-boundaries. However, when one considers the effects of head frequency, at least two interpretations are possible. We can either assume that the lexical unit corresponding to the head has already been accessed when the compound typing is initiated and remains active at least until the production of the second immediate constituent is initiated, or that it is first accessed when the typing of the first immediate constituent is achieved and the second immediate constituent is initiated. As our experiment is not designed to bring about a decision between these competing views, neither hypothesis can be supported or rejected. However, there is — at least for spoken language production — compelling evidence that the lexical unit corresponding to the compound head is already accessed at an early stage of compound production. Bien et al. (2005) found that the frequency of the head noun affects the naming latencies of compounds as indicated by shorter latencies for compounds with a high head frequency compared to those with a low-head frequency.

Additionally, the data analysis reveals an effect of semantic transparency although there is also an interaction of compound frequency and semantic transparency. As can be observed in Figure 1 (showing estimated marginal means from an optimal model), this interaction is a result of the fact that the difference between the opaque and transparent conditions (regardless of head frequency) is present in low-frequency compounds only. On average, SM-IKIs of opaque low-frequency compounds were estimated to be 18 ms faster than SM-IKIs of transparent low-frequency compounds, whereas SM-IKIs of opaque high-frequency compounds were almost as fast as those of transparent high-frequency compounds (see Figure 1). These findings demonstrate that semantic transparency plays a different role in the production of low and high-frequency compounds. This disparity can be accounted for by assuming that, at the semantic level, transparent low-frequency compounds have a meaning representation that is more strongly connected to their constituent morphemes, whereas opaque low-frequency compounds have a meaning representation that is more strongly connected to the whole compound. High-frequency compounds, however, seem to have a meaning representation that is more strongly connected to the whole compound, regardless of its transparency/opacity. It is plausible to assume that the high frequency strengthens the connection with the whole compound representation at the semantic level. The effect of semantic transparency found in low-frequency compounds indicates that the meaning representation is active at SM-boundaries. According to the general view that whole-word access procedures tend to be faster than compositional access procedures due to the fact that they involve less computational steps (e.g. McQueen & Cutler, 1998), the shorter SM-IKIs of opaque compounds compared to those of transparent ones can be accounted for as follows: the semantic interpretation of opaque compounds relies more on the meaning representation of the whole compound, whereas the semantic interpretation of transparent compounds relies more on the meaning representations of the constituent morphemes. Hence, opaque compounds show a processing advantage compared to transparent compounds.

Although we found effects of semantic transparency, our findings suggest that both transparent and opaque compounds show morphological constituency as indicated by the constant effect of head frequency. These findings indicate that compositional procedures are always involved in the production of German compounds and that compound production is sensitive to the morphological structure, independent of the transparency/opacity of the compound's constituents. Our findings confirm Aronoff's (1994) assumption that morphology operates by itself, i.e. that rules may govern the structure of polymorphemic words without simultaneous recourse to their meaning. These findings can be explained by assuming that both transparent and opaque compounds are equally composed at the form level but that, at the semantic level, the opaque compounds (and also transparent compounds, if high frequency) are more strongly connected to the meaning of the whole word form. This account is principally in line with the model of compound recognition and interpretation developed by Libben (1998) which locates word form representations at the lexical level and meaning representations at the conceptual level. Accordingly, the disparity between opaque and transparent conditions for low-frequency compounds observed in the present study might result from the different connections between compounds at the lexical level and their meaning representations at the conceptual level. While opaque compounds are only connected with the meaning representation of the whole word form, transparent compounds are also connected with the meaning of their constituents. The present finding that the production processes are sensitive to morphological structure independent of semantic transparency/opacity of the compound's constituents confirms the results of previous studies dealing with the impact of semantic transparency on compound comprehension (Zwitserlood, 1994; Libben, Gibson, Yoon & Sandra, 2003) and on the production of complex words (Roelofs & Baayen, 2002). In a lexical decision task, Zwitserlood (1994) found that English compounds equally primed their morpheme constituents, regardless of their semantic transparency/opacity, but that, at the semantic level, transparent compounds were connected to their constituent morphemes, whereas opaque ones were not. These results are supported by a more recent study using a constituent priming paradigm (Libben et al., 2003) in which the authors found that the prior presentation of a compound constituent facilitated the recognition of English compounds. This was the case although the semantic transparency/opacity provided evidence for the sensitivity of compound processing to the morphological structure of opaque as well as transparent compounds in language comprehension. In an on-line preparation paradigm task, Roelofs & Baayen (2002) compared the preparation effect for transparent complex nouns in Dutch like <bijval> ('applause', literally 'additional fall'), opaque complex nouns like <bijrol> ('supporting role', literally 'additional role') and monomorphemic words sharing the same initial syllable <bij> like <bijbel> 'bible'. The authors found that, concerning the preparation effect, opaque complex words behave like transparent complex words rather than monomorphemic words as indicated by the fact that the size of the morphemic effect was almost identical for transparent and opaque conditions. The authors interpret their findings as evidence for the so-called morphological autonomy hypothesis whereby all morphologically complex words are also represented in terms of their constituent morphemes. They also consider their findings as evidence against the so-called semantic dependency hypothesis, whereby only semantically transparent complex words are composed.

The findings that delays at SM-boundaries were affected by compound frequency as well as by head frequency indicate that the production of German compounds is accomplished by the activation of whole-word units as well as by the activation of their constituents. These findings are in line with dual-route models suggesting compositional as well as holistic processing for polymorphemic words. The cumulative effects of compound frequency and head frequency on SM-IKIs, however, cannot be accounted for by versions of the dual-route model which suggest that certain words are processed by the holistic route and others by the morphemic route (e.g. the Augmented Addressed Morphology Model [AAM], Caramazza et al., 1988; Chialant & Caramazza, 1995). The AAM assumes that known polymorphemic words are processed via the whole word route while novel words are processed via the morphemic route. For compounds, one could assume that such a distinction can be made on the basis of lexical frequency with high-frequency compounds being processed via the whole word route and low-frequency compounds via the morphemic routes. Our results, however, do not confirm this prediction. Rather, they suggest that both routes are involved in the production of high-frequency as well as low-frequency compounds and consequently they challenge such versions of dual-route models.

More adequate versions of dual-route models are those postulating the parallel activation of both routes (e.g. the Morphological Race Model [MRM], Frauenfelder & Schreuder, 1991; the meta-model, Schreuder & Baayen, 1995; Baayen & Schreuder, 1999). The MRM is a dual-route model in which two independent routes operate in parallel: the holistic and the morphemic route. In this model, both routes compete and the faster route wins the race. In so far as one route always wins the race, the MRM also fails to account for our findings which suggest that both routes are involved in the production of German compounds. Our findings are more compatible with the meta-model (Schreuder & Baayen, 1995; Baayen & Schreuder, 1999) in which the two routes converge, i.e. both routes contribute to the processing of polymorphemic words. This contribution depends on lexical frequency and semantic transparency. As the two routes converge, there is no need to assume that only one route is successful in processing a polymorphemic word. Our findings that the compound and head frequency crucially affects the time course of compound typing at SM-boundaries indicate the involvement of both routes and therefore clearly support the meta-model assumption that the two routes converge. Furthermore, by assuming that the two routes operate in parallel without an a priori fixed order, the meta-model provides a relatively straightforward explanation for the finding that compound and head frequency affect the time course of compound typing at more or less the same point in production, namely at SM-boundaries. Based on the same assumption, the model can also account for the strong effects of compound frequency at this relatively late production stage: the compound representation remains active during the production process even when the production of the second immediate constituent, i.e. the head, is initiated. Moreover, compound representations seem to have a higher activation level compared with representations of the head as indicated by the hierarchy of SM-IKIs found in our experiment.

There are some parallels between our findings and the results of previous studies dealing with the production and comprehension of compounds. On the one hand, our findings confirm the results of the neurolinguistic studies outlined in the introduction (Delazer & Semenza, 1998; Blanken, 2000; Badecker, 2001) which state that compositional as well as holistic procedures are involved in the production of compounds. On the other hand, similar results concerning the effects of compound and head frequency, as well as their overlap in time, were found in a comprehension study on Finnish compounds. Pollatsek, Hyönä & Betram (1998) measured eye movements while sentences containing compounds were read silently. They found that compound and constituent frequency affect fixation durations at different stages of reading and that the frequency of the second constituent "affect[s] processes at more or less the same time as whole-word frequency" (Pollatsek, Hyönä & Betram 1998: 831). From these findings they infer that holistic procedures and compositional procedures occur in parallel.

When comparing our findings with the results of Bien et al. (2005), the only study dealing with frequency effects in compound production in normal adults, we recognise similarities as well as differences. Our study confirms the finding of Bien et al. (2005) that head frequency is an influencing variable in compound production. Nevertheless, in contrast to their findings that "compound frequency play[s] a minor role only" (p. 17881), we found a highly significant effect of compound frequency. This disparity might result from the fact that our measurements and those of Bien et al. (2005) took place at different stages of compound production: we assessed the influence of frequency by measuring the intervals between the last letter/key of the first constituent and the first letter/key of the second constituent of compounds, whereas Bien et al. (2005) measured speech onset latencies. It is likely that compound frequency affects later stages of the production process rather than earlier ones. The disparity could also be ascribed to the different language modalities examined. In the present study, we investigated written language production, whereas Bien et al. (2005) looked at spoken language production. Therefore, in order to clarify the impact of compound frequency on compound production, more studies are needed.

Our study does not confirm the findings of Dohmes, Zwitserlood & Bölte (2004) referred to in the introduction that semantic transparency has no impact on the production of compounds. This discrepancy could arise from two factors: 1) from the different experimental paradigms used in the studies, and 2) from the different language modalities examined: written language in the present study vs. spoken language in the Dohmes, Zwitserlood & Bölte (2004) study. In view of the fact that, apart from these two studies, to our knowledge there has been no further research into the impact of semantic transparency on compound production, further studies are needed to clarify the role of semantic transparency in the production of compounds.

To sum up, our findings clearly suggest that holistic as well as compositional procedures are involved in the written production of German compounds as indicated by the effects of compound frequency and head frequency at SM-boundaries. Furthermore, we found that opaque as well as transparent compounds show effects of head frequency providing evidence for morphological composition in compound production. This suggests that the processing system is sensitive to the morphological structure of compounds independent of their semantic transparency/opacity. The effect of semantic transparency, found with low-frequency compounds, can be accounted for by assuming that transparent compounds have a meaning representation that is more strongly connected to their constituents, whereas opaque ones have a meaning representation that is more strongly connected to the whole word form.

References

Aronoff, M. (1994). *Morphology by itself*. Cambridge, MA: MIT Pres.

Baayen, R.H., Piepenbrock, R. & van Rijn, H. (1993). *The CELEX lexical database [CD-ROM]*. Philadephia, P.A: Linguistic Data Consortium, University of Pennsylvania.

Baayen, H. & Schreuder, R. (1999). War and peace: Morphemes and full forms in a noninteractive activation parallel dual-route model. *Brain and Language* 68: 27–32.

Badecker, W. (2001). Lexical composition and the production of compounds: Evidence from errors in naming. *Language and Cognitive Processes* 16: 337–366.

Bien, H., Levelt, W.J.M. & Baayen, H.R. (2005). Frequency effects in compound production. *Proceedings of the National Academy of Sciences of the United States of America* 102(49): 17876–17881.

Blanken, G. (2000). The production of nominal compounds in aphasia. *Brain and Language* 74: 84–102.

Delazer, M & Semenza, C. (1998). The processing of compounds words: A study in aphasia. *Brain and Language* 61: 54–62.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (ed.), *Language Production*, vol. 2, 257–294. San Diego: Academic Press.

Caramazza, A.; Laudanna, A. & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition* 28: 297–332.

Chialant, D. & Caramazza, A. (1995). Where is morphology and how is it processed? The case of written word recognition. In L.B. Feldman (ed.), *Morphological Aspects of Language Processing*, 55–76. Hillsdale: Erlbaum.

Dohmes, P., Zwitserlood, P. & Bölte, J. (2004). The impact of semantic transparency of morphologically complex words on picture naming. *Brain and Language* 90: 203–212.

Frauenfelder, U.H. & Schreuder, R. (1991). Constraining psycholinguistic models of morphological processing and representation: the role of productivity. In G. Booij & J. van Marle (eds.), *Yearbook of Morphology*, 165–183. Dordrecht/Boston/London: Kluwer.

Jescheniak, J.D. & Levelt, W.J.M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Language, Memory and Cognition* 20: 824–843.

Largy, P., Fayol, M. & Lemaire, P. (1996). The homophone effect in written French: The case of verb-noun inflection errors. *Language and Cognitive Processes* 11: 217–255.

Libben, G. (1998). Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. *Brain and Language* 61: 30–44.

Libben, G.; Gibson, M.; Yoon, Y.B. & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language* 84: 50–64.

McQueen, J.M. & Cutler, A. (1998). Morphology in Word Recognition. In A. Spencer & A. M. Zwicky (eds.), *The Handbook of Morphology*, 406–427. Oxford: Blackwell.

Nottbusch, G., Grimm, A., Will, U. & Weingarten, R. (2005). Syllabic structures in typing: Evidence from deaf writers. *Reading and Writing* 18: 497–526.

Nottbusch, G., Weingarten, R. & Sahel, S. (2006). From written word to written sentence production. In D. Galbraith, L. van Waes & M. Torrance (eds.), *Writing and Cognition. Research and Applications* (*Studies in Writing* 4), 31–54. Amsterdam: Elsevier.

Orliaguet, J.-P. & Boë, L.-J. (1993). The role of linguistics in the speed of handwriting movement: Effects of spelling uncertainty. *Acta Psychologica* 82: 103–113.

Pollatsek, A., Hyönä, J. & Betram, R. (1998). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology* 26: 820–833.

Pynte, J., Courrieu, P. & Frenck, Ch. (1991). Evidence of repeated access to immediate verbal memory during handwriting. *European Journal of Psychology of Education* VI(2): 121–125.

Rapp, B., Epstein, C. & Tainturier, M.-J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive Neuropsychology* 19: 1–29.

Roelofs, A. & Baayen, H. (2002). Morphology by itself in planning the production of spoken words. *Psychonomic Bulletin & Review* 9(1): 132–138.

Sahel, S., Nottbusch, G., Blanken, G. & Weingarten, R. (2005). The role of phonology and syllabic structure in the time course of typing: Evidence from aphasia. *Linguistische Berichte* 201: 65–87.

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology* 42A, 529–567.

Sandra, D. (2003). Homophonous regular verb forms with a morphographic spelling: Spelling errors as a window on the mental lexicon and working memory. In E.M.H. Assink & D. Sandra (eds.), *Reading Complex Words. Cross-Language Studies*, 315–330. New York: Kluwer.

Schreuder, R & Baayen, H. (1995). Modeling Morphological Processing. In L.B. Feldman (ed.), *Morphological Aspects of Language Processing*, 131–154. Hillsdale: Erlbaum.

Taft, M. & Forster, K. (1975). Lexical storage and retrieval of prefixes words. *Journal of Verbal Learning and Verbal Behavior* 14: 638–647.

Weingarten, R., Nottbusch, G. & Will, U. (2004). Morphemes, syllables and graphemes in written word production. In Pechmann & Ch. Habel (eds.), *Multidisciplinary approaches to speech production*, 529–572. Berlin: Mouton de Gruyter.

Will, U., Nottbusch, G. & Weingarten, R. (2006). Linguistic units in word typing: Effects of word presentation modes and typing delay. *Written language and literacy* 9(1): 153–176.

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes* 9: 341–368.

Corresponding Author

Said Sahel
Faculty of Linguistics and Literary Studies
University of Bielefeld
PO Box 100131
33501 Bielefeld
Germany

said.sahel@uni-bielefeld.de