Frequency effects in compound production

Heidrun Bien*, Willem J. M. Levelt*, and R. Harald Baayen*

Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH, Nijmegen, The Netherlands

Contributed by Willem J. M. Levelt, October 7, 2005

Four experiments investigated the role of frequency information in compound production by independently varying the frequencies of the first and second constituent as well as the frequency of the compound itself. Pairs of Dutch noun–noun compounds were selected such that there was a maximal contrast for one frequency while matching the other two frequencies. In a position–response association task, participants first learned to associate a compound with a visually marked position on a computer screen. In the test phase, participants had to produce the associated compound in response to the appearance of the position mark, and we measured speech onset latencies. The compound production latencies varied significantly according to factorial contrasts in the frequencies of both constituting morphemes but not according to a factorial contrast in compound frequency, providing further evidence for decompositional models of speech production. In a stepwise regression analysis of the joint data of Experiments 1–4, however, compound frequency was a significant nonlinear predictor, with facilitation in the low-frequency range and a trend toward inhibition in the high-frequency range. Furthermore, a combination of structural measures of constituent frequencies and entropies explained significantly more variance than a strict decompositional model, including cumulative root frequency as the only measure of constituent frequency, suggesting a role for paradigmatic relations in the mental lexicon.

lexical access | morphology | speech production

High-frequency words are produced faster than low-frequency words. Since the seminal study of Oldfield and Wingfield (1), the effect of word frequency has emerged to be replicable and robust. In their series of timed picture-naming studies in seven languages, Bates *et al.* (2) found large frequency effects in all of the seven languages studied. Jescheniak and Levelt (3) observed that the frequency effect for lemma retrieval diminished quickly over repetition but that the frequency effect for a word's form (lexeme) remained stable across repetitions. The cumulative homophone effect reported in that study suggests that the effect of word frequency arises at the level of word form, rather than conceptualization or articulation. Word frequency has, therefore, been attributed to the access of a word's phonological code (ref. 4, but see refs. 5 and 6). The general finding that a word's frequency is correlated with its production latency has become a powerful experimental tool.

In this study, we address frequency effects in the production of Dutch compounds. Fully nondecompositional theories predict frequency effects for each individual form of occurrence. In such theories, only the specific frequency of morphologically complex words such as *handbag* is expected to be predictive of their production latency. We will refer to this form-specific frequency as the word form frequency. In the case of compounds such as *handbag*, fully nondecompositional theories distinguish the word form frequency of the singular *handbag* from the word form frequency of the plural *handbags*.

In a fully decompositional model, such as in refs. 4 and 7 and its computer simulation WEAVER++ (8), all complex words are assembled from their constituent morphemes. The more often a morpheme has been used, the lower its activation threshold. Hence, the proper frequency measure for predicting production latencies in WEAVER++ for words such as *hand* is the summed frequency of all variants of that word, whether part of inflected (*hands*), derived (*handy*) or compound (*handbag*) words. Each of those occurrences is assumed to leave a frequency trace on the stem *hand*. In what follows, we will refer to the summed frequencies of a word as its cumulative root frequency (see refs. 9–11 for cumulative root frequency effects reported for comprehension).

According to WEAVER++ (8), the cumulative stem frequencies of the constituents *hand* and *bag* should be the relevant frequency measures for predicting the production latency of a compound, such as *handbag*. Roelofs (12) addressed the question of whether the form lexicon underlying speech production contains morphologically decomposed entries by using the implicit priming paradigm (13, 14). In this paradigm, subjects produce words from learned paired-associates. Homogenous response sets, in which all response words began with the same morpheme, resulted in shorter naming latencies than heterogeneous response sets, in which all initial morphemes of response words were different. Crucially, this preparation effect was larger for words with initial low-frequency morphemes than for ones with high-frequency morphemes, and the effect was stable in repeated measurements because low-frequency morphemes have more to gain from implicit priming than high-frequency morphemes. This finding supports decompositional theories in which the constituents of words like *handbag* are individually accessed during the production process. Currently, WEAVER++ implements the most parsimonious decompositional theory by assuming that the form representation for *hand* in *handbag* is the same as the one for the word *hand* itself.

An intermediate position between nondecompositional and fully decompositional models is to assume structured storage. Instead of storing *handbag* at the form level as two independent monomorphemes, it might be stored with information about their combination. In a model with structured storage, the frequencies of *hand* as the first constituent of any compound or *bag* as the head of any compound can be more precise predictors than their frequencies as independent words. We will refer to a set of compounds sharing the same left (or right) constituent as the left (or right) constituent family, as in ref. 15. For each constituent family, we have a distribution of the compound frequencies of its members. One way to obtain a point estimator of such a distribution is to sum the frequencies of its members (henceforth, positional frequency). Another point estimator of this distribution is to compute Shannon's entropy for the probability distribution estimated by the relative frequencies (the frequencies normalized with positional frequency, henceforth, positional entropy) (16, 17). Both the positional frequency and the positional entropy are measures calculated for the range of alternative compounds sharing the same morpheme in the same position. Whereas the positional frequency adds up the frequencies of the constituent family members, the positional entropy takes into account the probability distribution within the family. Other constituent frequency measures that are of potential interest are the summed frequencies of all other complex words in which the constituent appears (henceforth, complement frequency) and the entropy of the constituents calculated over the full range of

Conflict of interest statement: No conflicts declared.

Freely available online through the PNAS open access option.

*To whom correspondence may be addressed. E-mail: heidrun.bien@mpi.nl, pim@mpi.nl, or baayen@mpi.nl

morphologically complex words in which the constituent appears (henceforth, derivational entropy). Finally, we define the summed frequencies of a word's inflectional variants as lemma frequency. In the case of a simple or complex noun, the inflectional variants are the singular and the plural forms. The constituent lemma frequency of *hand* in *handbag*, therefore, refers to the sum of the frequencies of *hand* and *hands*, whereas the lemma frequency of the compound (henceforth, compound frequency) refers to the sum of the frequencies of *handbag* and *handbags*, respectively. In general, the left (or right) constituent lemma frequency, cumulative root frequency, positional frequency, positional entropy, complement frequency, and derivational entropy are strongly correlated.

Key Questions

In this paper we address three key questions. First, can separate frequency effects for the constituents of compounds be ascertained? Constituent frequency effects would provide further evidence against full unstructured storage models. Roelofs (12) observed constituent frequency effects for the left constituent by using implicit priming. We seek to replicate this finding by using immediate naming. In addition, we examine whether a similar frequency effect can be found for the right constituent. It is not self-evident that a frequency effect for the right constituent should exist, even within decompositional theories. Various studies (12, 18) have shown that production proceeds incrementally, suggesting that the frequency of the second constituent might become relevant only after completion of the planning and initiation of the articulation of the first constituent. However, there are circumstances in which the length of the word codetermines object-naming latencies (19), indicating that speakers may plan the complete phonological word before speech onset. This evidence is in line with recent studies addressing the acoustic realization of complex words. Stems pronounced in isolation tend to have longer durations and tend to be produced with a different intonation contour than the same stems appearing as the initial constituents of complex words, both for inflection and derivation (20–22). This finding suggests that the planning of the articulation of the first constituent is to some extent dependent on the presence of a second constituent.

Second, does compound frequency contribute to response latency? In nondecompositional models, compound frequency should be the only relevant measure, but in strict decompositional models it should be irrelevant. In a model with structured storage, a compound frequency effect cannot be ruled out, but it might be strongly modulated by the role of the constituent families.

Third, are constituent frequency effects on response latency best predicted from the cumulative root frequencies of the constituents, or do we rather see different effects of the more specific frequency measures, including positional measures based on the constituent families? This comparison allows us to distinguish between full decomposition as in WEAVER++ and structured storage.

In our experiments, we systematically manipulated the frequency of the first and second morphemes and the compound itself to examine their individual influences on compound production latency. For each experiment, we selected pairs of Dutch compounds as targets such that there was a maximal contrast for one frequency factor, whereas the other two (except in Exp. 3) were matched. The contrasts were constructed in terms of lemma frequency but coincided with a range of additional contrasts, such as cumulative root frequency, positional frequency, or derivational entropy for the constituents. After the factorial analyses of Exps. 1–4 we present a regression analysis addressing the question, which of these frequency measures are the appropriate predictors?

All compounds used in this study are semantically transparent Dutch noun–noun compounds, with the first constituent being the modifier and the second constituent being the head noun.

Materials and Methods

Material Selection. From the CELEX lexical database (23), we selected Dutch noun–noun compounds on the basis of three frequency counts: the lemma frequency (the summed frequencies of the word's inflectional variants) of the compound, the lemma frequency of its left constituent, and the lemma frequency of its right constituent. For each experiment, we selected 16 pairs of compounds. All compounds are listed in Table 3, which is published as supporting information on the PNAS web site.

Pairs in Exp. 1: The Frequency of the Head Noun. In Exp. 1, the compounds in a pair shared the first morpheme (e.g., *luchtbrug–luchtbuks*/airlift–airgun). The pairs were matched for compound frequency and differed with respect to the frequency of the second constituent, which was either high (mean, 13,354; median, 5,731; range, 437–50,439) or low (mean, 145; median, 61; range, 1–867). (All CELEX frequencies reported here and below are counts based on a corpus of 42 million words.)

The 16 compounds with low-frequency second constituents were arranged into eight pairs (e.g., *luchtbuks–broodkruim*) with the constraint that the pair mates had minimal phonological overlap, no obvious semantic relation, and had similar compound frequencies. The same was done for the 16 compounds with high-frequency second constituents. Presentation was blocked by condition, and the order of the blocks was counterbalanced. Half of the participants started out with the eight pairs with low-frequency second constituents, and the other half began with the eight pairs with high-frequency heads. We blocked the conditions to minimize the effects of criterion-setting that would lead to elongated responses for otherwise short reaction times and to speeded responses for otherwise long reaction times (19, 24). The 16 subsets of target compounds were complemented by three practice subsets, with compounds of similar structure and frequency.

Pairs in Exp. 2: The Frequency of the Modifier. In Exp. 2, the compounds in a pair shared the second morpheme and were matched for compound frequency. The first morpheme carried a factorial contrast between high (mean, 8,072; median, 7,111; range, 1,424–23,062) and low (mean, 660; median, 356; range, 39–2,645) frequency. As described for Exp. 1, we blocked the conditions by rearranging the items into eight subsets of compounds with low-frequency modifiers and eight subsets of compounds with high-frequency modifiers and added three practice subsets of similar structure and frequency.

Pairs in Exp. 3: Frequency Contrasts for Head and Modifier. In Exp. 3, we selected 32 compounds pairwise-matched for compound frequency. Within a pair, one compound had high-frequency constituents (mean, 10,400; median, 7,213; range, 409–48,452), and the other had low-frequency constituents (mean, 618; median, 291; range, 4–4,416). A given constituent appeared in only one compound. We used the same blocking strategy as described for Exps. 1 and 2 and complemented the resulting 16 subsets of condition–intern repairings with three practice subsets of similar structure and frequency.

Pairs in Exp. 4: The Frequency of the Compound. In Exp. 4, 32 compounds were selected that were pairwise-matched according to the frequency of the first morpheme and according to the frequency of the second morpheme, whereas the frequency of the compound was either high (mean, 973; median, 897; range, 516–2,369) or low (mean, 48; median, 39; range, 6–132). As described for the previous experiments, we blocked the conditions, creating eight subsets of low-frequency compounds and eight subsets of high-frequency compounds that were complemented by three similar practice subsets.

Participants. For each experiment, 24 native speakers of Dutch were recruited from the subject pool of the Max Planck Institute for Psycholinguistics. None of the subjects took part in more than one of the experiments. Each participant received 5.00 € for participating.

Position–Response Association Task. Participants were tested individually in a dimly lit, sound-attenuated booth. Participants were comfortably seated in front of a cathode ray tube computer screen, a Sennheiser microphone, and a cordless mouse, and they wore headphones. On average, a session lasted 45 min.

We used a position–response association task (25), in which participants learned to associate the two compounds in a subset with visually marked positions on the left and right part of a computer screen. For each subset, the experimental procedure consisted of a learning phase, a practice phase, and a test phase. Each phase was introduced by an attention signal presented on the screen for 2 s and ended with a pause signal that remained on the screen until the following phase was initiated by the experimenter.

In the learning phase, participants were presented with the two compounds over headphones. Simultaneously with hearing the first compound, they saw the icon of a loudspeaker appearing on the left side of the screen. Simultaneously with hearing the second compound, the same icon appeared on the right side of the screen. This procedure was repeated once.

In the practice phase, both icons (left and right) were visible with the cursor of the mouse in the center of the screen. The subject was then acoustically presented with one of the two compounds and had to click on the associated icon. Both compounds were presented twice and in random order. We provided the participants with feedback on their accuracy by displaying the number of errors (0–4) on the screen.

In the test phase, 20 trials of a distractor task alternated with 20 trials of the experimental task. A test phase always began with a distractor trial. In a distractor trial, one of four single-digit numbers (1, 2, 3, or 6) was presented in the center of the screen and had to be named as fast and correctly as possible. We included those distractor trials to avoid that participants would have to produce exactly the same word during consecutive trials. In other words, the insertion of distractor trials made it difficult for participants to use the break between two experimental trials to already prepare one of the target words. In an experimental trial, the icon of a loudspeaker was presented either on the left or right position of the screen, prompting the participant to say aloud the associated compound again as fast and correctly as possible. Each position appeared a total of 10 times. The two positions were presented in pseudo random order with the restriction of a maximum of four consecutive repetitions of one position.

Simultaneous with the presentation of the icon, the voice key was activated for 1,500 ms. Naming latencies longer than 1,500 ms were counted as time-outs. The experimenter monitored the participant's responses through headphones and took notes of incorrect naming, hesitations, and voice key errors.

Results

Only those compounds for which a correct response was obtained were included in the analysis. Time-out trials (>1,500 ms) and extreme outliers [i.e., latencies outside a range of two standard deviations around the mean latency for each subject per condition (high or low) as well as for each item] were also removed from the analysis.

Exp. 1: The Frequency of the Head Noun. Altogether, 421 trials were excluded (5%) in Exp. 1. Mean latencies, standard deviations, and error rates are summarized in Table 1.

We analyzed the latencies by subjects and by items, with frequency as a within factor, and order as a between factor. Compounds with a high right-constituent frequency elicited shorter latencies (on average 14 ms) than compounds with a low-frequency head, both in the analysis by participants [$F1(1, 22) = 5.8, P = 0.025$] and in the analysis by items [$F2(1, 30) = 18.1, P < 0.001$]. There was no effect of order in the by-participant analysis [$F(1, 22) = 2.1, P = 0.16$] and no interaction of order by frequency ($F(1, 22) = 1.8, P = 0.19$). In the by-item analysis, order emerged as a significant main effect [$F(1, 30) = 16.9, P < 0.001$] in interaction with frequency [$F(1, 30) = 5.9, P = 0.021$]. The interaction points to a significant difference between the 6-ms frequency effect for the high–low block order and the 21-ms frequency effect for the low–high block order. An analysis of the error scores revealed no significant main effects nor any interactions.

Table 1. Mean latencies for Exps. 1–4

Exp.	Frequency	Mean, ms (%)	LH, ms	HL, ms
1	High	457 ± 111 (3)	437	476
	Low	471 ± 116 (2)	458	482
2	High	443 ± 118 (5)	439	447
	Low	468 ± 129 (5)	487	450
3	High	414 ± 105 (6)	405	424
	Low	441 ± 115 (5)	445	437
4	High	442 ± 108 (4)	430	454
	Low	434 ± 104 (4)	433	435

Values are for the main effect of frequency ± standard deviation (with error percentages in parentheses) and for the block orders low–high (LH) and high–low (HL).

Apparently, the blocking strategy, which was chosen to avoid criterion setting, created an alternative problem: the interaction of frequency and practice. Because of practice, participants became faster, leading to shorter latencies in the second block compared with the first block. The speeding up was strong in the otherwise slow low-frequency set, whereas the already fast production of high-frequency items could hardly benefit from an additional effect of practice, underestimating the difference between low- and high-frequency items in the high–low order of presentation.

In short, the frequency of the head noun codetermines production latencies, although it is not the initial constituent of a compound, providing evidence for decompositionality as well as evidence against strict incrementality. Apparently, articulation is not initiated before the phonological code of the head noun has been retrieved. Exp. 2 investigates the predictivity of the frequency of the initial constituent.

Exp. 2: The Frequency of the Modifier. Altogether, 814 trials (10%) were removed from the analysis by following the criteria described above.

Compounds with a high left-constituent frequency elicited shorter latencies (on average 25 ms) than compounds with a low-frequency modifier in the analysis by participants [$F1(1, 22) = 19.4, P < 0.001$] and in the analysis by items [$F2(1, 30) = 48.5, P < 0.001$]. In the by-participant analysis, there was no main effect of order [$F(1, 22) = 0.2, P = 0.675$] but an interaction of order by frequency [$F(1, 22) = 13.5, P = 0.001$]. In the by-item analysis, order emerged as a significant main effect [$F(1, 30) = 9.0, P = 0.005$] in interaction with frequency [$F(1, 30) = 37.7, P < 0.001$]. As in Exp. 1, the difference between the high- and low-frequency conditions was bigger in the low–high block order (48 ms) than in its reverse (3 ms). An error analysis revealed no significant main effects nor interactions.

In short, this experiment shows that the frequency of the initial constituent affects compound production, as expected in a decompositional theory of compound production. We next investigated whether constituent frequency effects can be observed in the absence of compounds in the experiment that share head or

Table 2. The frequency characteristics of the material in Exps. 1–3 (log-transformed) for the left/right constituent

Exp.	Freq.	CumFreq.	LemFreq.	ComplFreq.	PosFreq.	PosEntr.	DerEntr.
1	High	9.44/10.01	8.89/9.59	8.59/8.95	8.00/7.89	13.87/10.47	21.39/14.78
	Low	9.44/6.88	8.89/5.76	8.59/6.49	8.00/5.91	13.87/2.46	21.39/9.36
2	High	9.61/9.59	9.00/9.01	8.73/8.59	7.90/7.99	12.30/12.97	22.98/16.86
	Low	7.29/9.59	6.50/9.01	6.64/8.59	6.28/7.99	7.27/12.97	9.31/16.86
3	High	9.99/9.82	9.43/9.03	9.04/8.99	7.77/7.90	13.12/11.32	22.80/18.24
	Low	7.25/7.86	6.42/6.44	6.65/6.95	5.68/6.64	6.00/6.17	12.79/12.26

Values are shown in ms. CumFreq., cumulative frequency; LemFreq., lemma frequency; ComplFreq., complement frequency; PosFreq., positional frequency; PosEntr., positional entropy; DerEntr., derivational entropy.

modifier. To maximize constituent effects, only two conditions were tested: high versus low frequency for both constituents.

Exp. 3: Frequency Contrasts for Head and Modifier. Time-out trials, voice key errors, extreme outliers, and incorrect naming responses were removed from the data set (863 trials, 11%).

Analyses of variance with frequency as a within factor and order as a between factor revealed that compounds with high-frequency constituents elicited shorter latencies (on average 27 ms) than compounds with low-frequency constituents [$F1(1, 22) = 20.7, P < 0.001$; $F2(1, 30) = 42.7, P < 0.001$]. There was no main effect of order [$F1(1, 22) = 0.2, P = 0.631$; $F2(1, 30) = 1.4, P = 0.241$], but an interaction of order by frequency [$F1(1, 22) = 8.5, P = 0.008$; $F2(1, 30) = 11.1, P = 0.002$]. As before, the interaction suggested that the difference between the high- and low-frequency conditions was more prominent in the low–high block order (40 ms) than in the high–low block order (13 ms). Analysis of the error scores revealed no significant main effects nor any interactions.

This experiment provides further support for the constituent frequency effects of Exps. 1 and 2, although not differentiating between the frequency effects of the first and second morphemes. Exp. 3 also rules out the possibility of a confound due to prior experience with a head or modifier in the experiment. In our final experiment, we addressed the question of whether the production latency of a compound might be additionally affected by the compound's own frequency of occurrence.

Exp. 4: The Frequency of the Compound. A total of 652 trials (8%) was excluded from the analyses by following the criteria defined above.

We analyzed the latencies with analyses of variance by participants and by items. Frequency was a within factor, and order was a between factor. On average, the high-frequency compounds elicited latencies that were 8 ms longer than the low-frequency compounds. This difference in the direction of an antifrequency effect did not reach full significance in the by-participants analysis [$F1(1, 22) = 3.5, P = 0.074$] but did reach significance in the by-item analysis [$F2(1, 30) = 6.8, P = 0.014$]. In the analysis by participants, there was no main effect of order [$F(1, 22) = 0.4, P = 0.515$] but an interaction of order by frequency [$F(1, 22) = 5.6, P = 0.028$]. In the analysis by items, order emerged as a significant main effect [$F(1, 30) = 9.9, P = 0.004$] in interaction with frequency [$F(1, 30) = 11.7, P = 0.002$]. The interaction points to a significant difference between the 19-ms antifrequency effect for the high–low block order and the 3-ms frequency effect for the low–high order. We have argued so far that the interaction of block order and frequency reflects a practice effect, which leads to an underestimation of the frequency effect when the slow block is presented last. In Exps. 1–3, the slower block clearly was the block with the items of the low-frequency condition. Here, the situation seems to be reversed, with the items of high compound frequency displaying a practice effect. If so, the antifrequency effect would be underestimated in the low–high block order. An analysis of the error scores revealed no effects.

In summary, high-frequency compounds were not produced any faster than low-frequency compounds. If anything, high-frequency compounds elicited longer naming latencies.

Comparing Frequency Measures. Material selection for Exps. 1–4 was based on lemma frequency. Table 2 shows for Exps. 1–3 that the high- versus low-frequency conditions for the constituents implemented contrasts in all of the different measures of constituent frequency and entropy that we defined in the introduction.

To further examine the predictivity of these measures, we included them in a stepwise regression analysis of the joint data of Exps. 1–4 along with neighborhood density (26), defined as the number of words that are similar to a target on the basis of the substitution of a single phoneme only (27), and factors controlling for the sensitivity of the voice key, addressing the nature of the onset phoneme. We started out with a variety of specifications, such as voicing, fricatives, and nasality, but only the factor plosive vs. nonplosive turned out to be a significant predictor. Fig. 1 visualizes the partial effects of the covariates.

A stepwise multilevel analysis of covariance (28–31), with participant as main grouping factor, revealed effects of the manner of articulation of the initial consonant [$\hat{\beta} = 0.0240, t(28822) = 7.5595, P < 0.0001$], and plosives elicited longer naming latencies (Fig. 1*I*), probably an artifact of the voice key. Fig. 1 *A* and *D* illustrates the facilitatory, linear effects of the left and right positional entropies adjusted for the effects of the other covariates [left positional entropy: $\hat{\beta} = -0.0081, t(28822) = -4.1796, P < 0.0001$; right positional entropy: $\hat{\beta} = -0.0098, t(28822) = -4.1716, P < 0.0001$]. Fig. 1 *B* and *E* pictures the facilitatory, linear effect of the left complement frequency [$\hat{\beta} = -0.0077, t(28822) = -6.2234, P < 0.0001$] and of the inhibitory, linear effect of the right complement frequency [$\hat{\beta} = 0.0055, t(28822) = 3.7080, P = 0.0002$]. Notice that this effect of complement frequency is significant for head and modifier but in the opposite direction. The more often the modifier appears as a constituent in other complex words independent from its position within those words, the faster the compound is named. In contrast, the more often the head constituent appears within other complex words, the slower the compound is named. For the head and only for the head, however, we also observe an effect of lemma frequency, and this effect is facilitative [$\hat{\beta} = -0.0036, t(28822) = -2.9686, P = 0.0030$]. The higher the frequency of the head as an independent word, the faster the compound is named (Fig. 1*F*). For the modifier, we further observe a linear, facilitatory effect of derivational entropy [$\hat{\beta} = -0.0071, t(28822) = -3.8231, P = 0.0001$] as plotted in Fig. 1*C*. Fig. 1*G* shows the nonlinear curve for the neighborhood density of the initial constituent [linear component: $\hat{\beta} = 0.0041, t(28822) = 5.2589, P < 0.0001$; quadratic component: $\hat{\beta} = -0.0001, t(28822) = -4.9338, P < 0.0001$], suggesting facilitation for left constituents with very sparse or very dense phonological similarity neighborhoods. There was no effect of the neighborhood density of the right constituent. Fig. 1*H* illustrates the nonlinear curve for compound frequency [linear

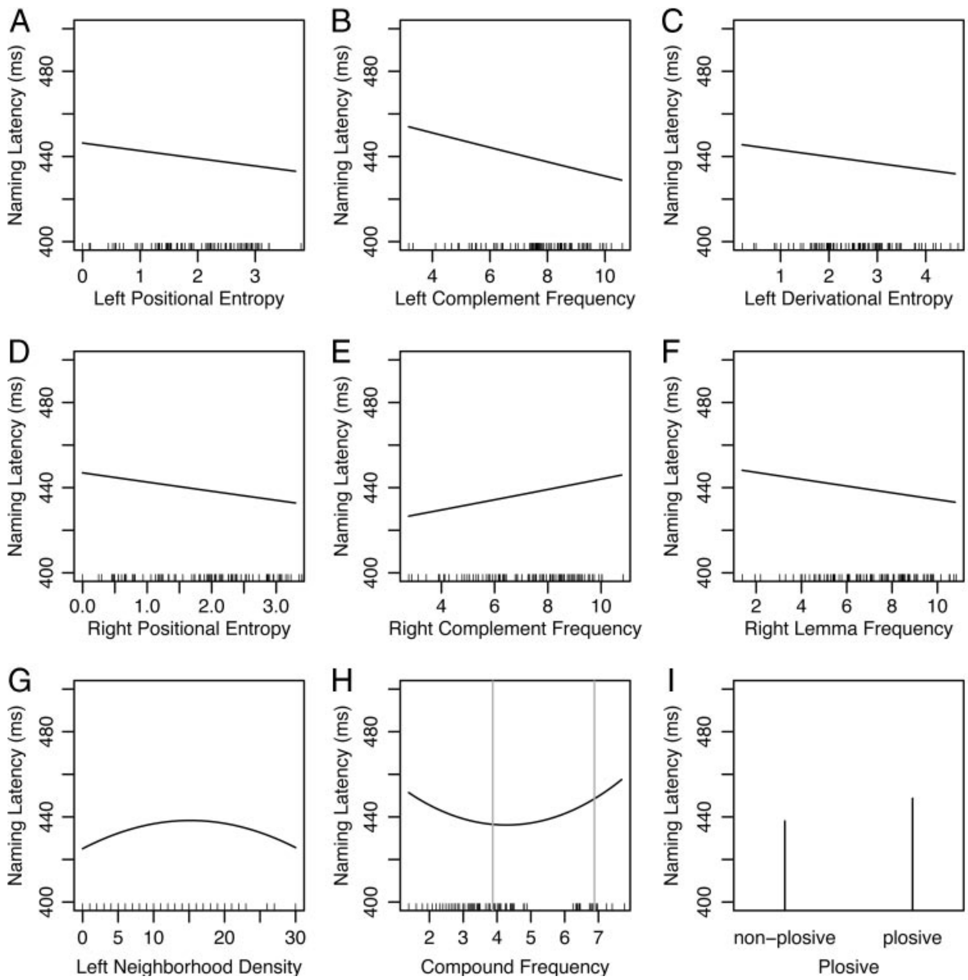

Fig. 1. Partial effects of the predictors in the multilevel covariance analysis of the data of Exps. 1–4. The left vertical axis shows the effect in log units; the right axis shows the effect in milliseconds. Values pertain to words that do not begin with a plosive and are adjusted for the effects of the other covariates at their median value.

component: $\hat{\beta} = -0.0351$, $t(28822) = -6.3141$, $P < 0.0001$; quadratic component: $\hat{\beta} = 0.0041$, $t(28822) = 6.5850$, $P < 0.0001$]. Although in the lower range of compound frequencies we see a facilitatory effect, this effect levels off and turns into inhibition in the higher range of compound frequencies. The two gray, vertical lines in Fig. 1*H* mark the averages of the low- and high-frequency conditions used in Exp. 3, illustrating why we did not observe a reliable effect of compound frequency in that experiment. The factorial contrast tended to balance low-frequency facilitation and high-frequency inhibition.

Finally, we compared the predictivity of our model with the predictivity of a strict decompositional model, a model in which the cumulative root frequencies of the left and right constituents were the only measures of constituent frequency entered into the regression equation. Both models included the nonfrequency predictors neighborhood density and voice key. The more complex model explained significantly more variance than the strict decompositional model ($P < 0.001$, log-likelihood ratio test) with a 61% increase in the variance explained by linguistic predictors.

In summary, the factorial analyses of Exps. 1–4 showed that the naming latency of a compound was affected by the frequencies of its constituents. The regression analysis revealed that the naming latency of a compound is affected by a combination of different measures of frequency and entropy for both constituents. Interestingly, a qualitative difference emerged with respect to how the left and right constituents were affected. Although for the modifier all significant effects of frequencies and entropies are facilitative, there is facilitation as well as inhibition for the head constituent. The total outcome is facilitation in both cases, with greater facilitation for the modifier. In the regression analysis, we also observed that compound frequency was one of the factors with explanatory value. The nonlinear effect of compound frequency suggests facilitation within the lower frequency range, combined with inhibition in the higher frequency range. This inhibition might represent a floor effect, however, because it might be an artifact of modeling nonlinearity with a simple, quadratic polynomial.

Discussion

This study addressed three key questions concerning the role of frequency in compound production. First, are there separate frequency effects for the constituents of a compound? Second, does the frequency of the compound itself affect its naming latency? Third, if we find effects of constituent frequency, which measures of frequency and entropy are the best predictors for the compound production latencies?

Exps. 1–3 addressed the first question by means of factorial contrasts. For pairs of compounds matched for compound frequency and sharing one constituent, a frequency contrast on the other constituent affected the production latencies. Both for the head (Exp. 1) and for the modifier (Exp. 2), a higher constituent frequency led to shorter naming latencies. This advantage of high-frequency constituents persisted in Exp. 3, in which both constituents were of high or low frequency. These results allowed us to conclude that the frequencies of both constituents indeed codetermine the production latency of a compound.

Exp. 4 addressed the second question by means of a factorial contrast, matching for constituent frequencies and contrasting compound frequencies. High-frequency compounds were not produced any faster than their matched counterparts with low frequencies. In fact, there was an indication that a high compound frequency might be inhibitory, but this inhibitory effect was small and not fully reliable.

To ascertain which frequency or entropy measures are the optimal predictors for the naming latencies, we analyzed the joint data of Exps. 1–4 by means of a multilevel regression analysis, which revealed that the production latencies were best predicted not by the constituent's cumulative frequency but rather by a combination of different, partly position-specific frequency and entropy measures and compound frequency.

These results shed new light on the role of decompositionality and incrementality in production. If compounds were similar to monomorphemic words, as in full-listing models, their naming latencies should depend on compound frequency only. Our experiments show, however, that compound frequency plays a minor role only, leading to facilitation only for the lowest ranges of compound frequencies and possibly to inhibition for the higher frequency ranges. The presence of an effect of compound frequency in the regression analysis shows that the position-response association task is, in fact, sensitive to word frequency, which has been demonstrated before (32). Because the effects that we observed for constituent frequencies were larger and more robust, we conclude that our data challenge models with only unstructured storage and no decomposition for complex words (see refs. 33 and 34 for similar conclusions based on aphasic speakers).

The constituent frequency effects observed for the left constituent replicate the frequency effect reported for initial constituents in ref. 12. The frequency effects observed for the right constituent provide further support for the possibility that production latencies are determined not only by the first morpheme or syllable but also by subsequent parts of the word as mentioned in ref. 19. The observation that frequency effects for the first constituent are more facilitatory than for the second constituent supports theories of incremental morphological processing in production. However, the effect of the second constituent argues against full incrementality. Speakers apparently plan the articulation of the first constituent with an eye on what is to be produced next. This look ahead may also shape the details of acoustic realization (20, 21).

The finding that the general frequency effects of the left and right constituents can be made more precise in terms of structural measures of constituent frequency and entropy offers new insights into the details of morphological processing in lexical access that invite further theoretical reflection. Our data suggest that the mental lexicon is highly sensitive to the specific morphophonological context in which a word has to be articulated.

The cumulative root frequency is a context-independent predictor of the speaker's familiarity with a given word form (e.g., *hand*), whereas position-specific measures are contextually conditioned predictors (e.g., *hand* in *handbag* or *handcuff*). This contextual sensitivity may well reflect the differences in the phonetic details of the production of *hand* by itself versus the production of *hand* as a head or a modifier. The positional frequency effects are in line with the predictions of decompositional models with structured storage.

The positional entropy effects provide further evidence for the role of paradigmatic relations (the links between morphologically related words) in the mental lexicon (15, 35). Paradigmatic effects in lexical processing show that words are not isolated processing units but rather structured units participating in networks of morphological relations. For instance, our positional entropy effects argue for structured storage, because the more often constituents appear in other compounds in the same structural position, the faster are their production latencies in immediate naming.

The similarity in the magnitude of the positional entropy effects for the left and right constituents suggests that the paradigmatic effects do not differentiate between the constituent that has to be pronounced first and the constituent that has to be pronounced last. However, from the perspective of incremental processing, simultaneous activation of the head with the modifier should be disadvantageous. In fact, there is evidence for some disadvantage associated with coactivation of the head: the inhibitory effect of the right-complement frequency. For the initial constituent, the modifier, all measures of frequency and entropy are facilitatory, but for the final constituent, the head, the inhibitory effect of the right-complement frequency modulates the facilitatory effect of the other facilitatory measures. Apparently, selecting the target's first constituent is harder the more other morphologically complex, noncompound words include the head constituent. Consequently, the overall frequency effect for the modifier emerged as stronger than the overall frequency effect for the head.

Considered jointly, our experiments support decompositional models of speech production in which the paradigmatic relations entertained by the constituents of a compound and the structural position of those constituents across the lexicon codetermine the details of the planning and articulation of the compound.

1. Oldfield, R. C & Wingfield, A. (1965) *Q. J. Exp. Psychol.* **17,** 273–281.
2. Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., Herron, D., Ching Lu, C., Pechmann, T., Pl′eh, C., *et al.* (2003) *Psychonom. Bull. Rev.* **10,** 344–380.
3. Jescheniak, J. D. & Levelt, W. J. M. (1994) *J. Exp. Psychol. Learn. Mem. Cognit.* **20,** 824–843.
4. Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999) *Behav. Brain Sci.* **22,** 1–38.
5. Caramazza, A., Costa, A., Miozzo, M. & Bi, Y. (2001) *J. Exp. Psychol. Learn. Mem. Cognit.* **27,** 1430–1450.
6. Jescheniak, J. D., Meyer, A. & Levelt, W. J. M. (2003) *J. Exp. Psychol. Learn. Mem. Cognit.* **29,** 432–438.
7. Levelt, W. J. M. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 13464–13471.
8. Roelofs, A. (1997) *Cognit.* **64,** 249–284.
9. Laudanna, A. & Burani, C. (1985) *Linguistics* **23,** 775–792.
10. Burani, C. & Caramazza, A. (1987) *Lang. Cognit. Processes* **2,** 217–227.
11. Schreuder, R. & Baayen, R. H. (1997) *J. Mem. Lang.* **37,** 118–139.
12. Roelofs, A. (1996) in *Yearbook of Morphology*, eds. Booij, G & van Marle, J. (Kluwer, Dordrecht, The Netherlands), pp. 135–154.
13. Meyer, A. S. (1990) *J. Mem. Lang.* **29,** 524–545.
14. Meyer, A. S. (1991) *J. Mem. Lang.* **30,** 69–89.
15. Krott, A, Baayen, R. H. & Schreuder, R. (2001) *Linguistics* **39,** 51–93.
16. Shannon, C. E. (1948) *Bell Syst. Tech. J.* **27,** 379–423.
17. Shannon, C. E. & Weaver, W. (1949) *The Mathematical Theory of Communication* (Univ. Illinois Press, Urbana).
18. Cholin, J., Schiller, N. O. & Levelt, W. J. M. (2004) *J. Mem. Lang.* **20,** 47–61.
19. Meyer, A. S., Roelofs, A. & Levelt, W. J. M. (2003) *J. Mem. Lang.* **48,** 131–147.
20. Kemps, R., Ernestus, M., Schreuder, R. & Baayen, R. H. (2005) *Mem. Cognit.* **33,** 430–446.
21. Kemps, R., Wurm, L., Ernestus, M., Schreuder, R. & Baayen, R. H. (2005) *Lang. Cognit. Processes* **20,** 43–73.
22. Koester, D., Gunter, T. C., Wagner, S. & Friederici, A. D. (2004) *J. Cognit. Neurosci.* **16,** 1647–1668.
23. Baayen, R. H., Piepenbrock, R. & Gulikers, L. (1995) The CELEX Lexical Database [CD-ROM] (Univ. Pennsylvania, Philadelphia), Release 2.
24. Lupker, S. J., Brown, P. & Colombo, L. (1997) *J. Exp. Psychol. Learn. Mem. Cognit.* **23,** 570–590.
25. Cholin, J., Levelt, W. & Schiller, N. O. (2005) *Cognition*, in press.
26. Vitevitch, M. S. (2002) *J. Exp. Psychol. Learn. Mem. Cognit.* **28,** 735–747.
27. Coltheart, M., Davelaar, E., Jonasson, J. T. & Besner, D. (1977) in *Attention and Performance*, ed. Dornick, S. (Erlbaum, Hillsdale, NJ), Vol. VI, pp. 535–556.
28. Pinheiro, J. C. & Bates, D. M. (2000) *Mixed-Effects Models in* S *and* S-PLUS, Statistics and Computing (Springer, New York).
29. Baayen, R. H., Tweedie, F. J. & Schreuder, R. (2002) *Brain Lang.* **81,** 55–65.
30. Baayen, R. H. (2004) *Ment. Lexicon Working Pap.* **1,** 1–45.
31. Quen′e, H. & Van den Bergh, H. (2004) *Speech Commun.* **43,** 103–121.
32. Cholin, J. (2004) Ph.D. thesis (Univ. Nijmegen, Nijmegen, The Netherlands).
33. Ayala, J. & Martin, N. (2002) *Brain Lang.* **83,** 81–83.
34. Mondini, S., Luzzatti, C., Saletta, P., Allamano, N. & Semenza, C. (2005) *Brain Lang.* **94,** 178–187.
35. Ernestus, M. & Baayen, R. H. (2004) *Linguistics* **42,** 873–903.