WORD AND SYLLABLE MODELS FOR GERMAN TEXT-TO-SPEECH SYNTHESIS

Bernd Möbius

Bell Labs – Lucent Technologies, Murray Hill, NJ 07974, USA

ABSTRACT

The correct pronunciation of unknown or novel words is one of the biggest challenges for text-to-speech systems. In this paper we describe the implementation of unknown word analysis as a central component of the text analysis module in the Bell Labs German text-to-speech system. The implementation is based on a model of the morphological structure of words and on the study of the productivity of word forming affixes. One important subcomponent of the word model is a phonotactic syllable model which enables the system to handle orthographic substrings that are unaccounted for by the explicitly listed morphemes. Finally, we discuss issues for future research.

1. INTRODUCTION

The correct pronunciation of unknown or novel words is one of the biggest challenges for text-to-speech (TTS) systems. Generally speaking, correct analysis and pronunciation can only be guaranteed if a direct match exists between an input string and a corresponding entry in a pronunciation dictionary or morphologically annotated lexicon. However, any well-formed text input to a general-purpose TTS system in any language is extremely likely to contain words that are not explicitly listed in the lexicon.

There are two major reasons why this statement holds. First, the inventory of lexicon entries of every language is unbounded; every natural language has productive word formation processes, and the community of speakers of a particular language can, via convention or decree, create new words as need arises. Second, the set of (personal, place, brand, etc.) names is, from a practical point of view, too large to be listed and transcribed in a pronunciation dictionary. Even more importantly, names are subjected to the same productive and innovative processes as regular words are. Thus, in unlimited vocabulary scenarios we are not facing a memory or storage problem but the requirement for the TTS system to correctly analyze unseen orthographic strings.

The German language is notorious for its extensive use of compounds. What makes this a challenge for linguistic analysis is the fact that compounding is extraordinarily productive. Whereas the famous *Donaudampfschiffahrtsgesellschaftskapitän* "captain of the steam boat shipping company on the Danube river" has almost achieved lexical status by now, examples of spontaneous novel word coinage by compounding are easy to come by: *Unerfindlichkeitsunterstellung* "allegation of incomprehensibility", for the purpose of this paper, by its author; or *Oberweserdampfschiffahrtsgesellschaftskapitänsmützenberatungsteekränzchen* "tea klatsch for the advice on yachting caps worn by captains of the steam boat shipping company on the upper Weser river", submitted by an anonymous user of our interactive TTS web site. Linguistic analysis has to provide a mechanism to appropriately decompose compounds and, more generally, to handle unknown words.

Arguably, names are a special category of unknown words. The analysis and pronunciation of names has been described in detail elsewhere [5]; we will therefore only briefly touch on this particular problem.

The unknown word analysis component of the German TTS system [9] is based on a model of the morphological structure of words. We also performed a study of the productivity of word forming affixes, applying the productivity measure suggested by Baayen [1]. This linguistic description was compiled into a weighted finite-state transducer (WFST). Finite-state transducer (FST) technology enables the dynamic and recursive combination of lexical and morphological substrings, which cannot be achieved by a static pronunciation dictionary. WFST technology is the framework for the text analysis components of the Bell Labs multilingual TTS system [13].

2. PRODUCTIVE WORD FORMATION

2.1. Productivity Measure

Morphological productivity has been defined as "the possibility for language users to coin, unintentionally, a number of formations which are in principle uncountable" [12] (cited in [8, p. 3]). Truly productive word formation processes are unintentional; for instance, a native speaker of German would not be aware of the fact that in our example *Unerfindlichkeitsunterstellung*, the first component *Unerfindlichkeit* results from a morphological process that transforms an adjective into a noun by appending a productive noun forming suffix. The suffix *-keit* is the default choice if an adjective ending on *-lich* is to be turned into a noun.

The ability to consciously coin new words by applying otherwise unproductive patterns has been refered to as "morphological creativity" [12]. The resulting novel word typically draws the attention of the listener or reader. The process of purposefully creating new words has been characterized as marginally productive. However, for the specific purpose of novel word analysis in TTS, the distinction between morphological creativity and productivity is largely irrelevant. It is more appropriate to consider these two notions as degrees or variants of the same linguistic phenomenon. What is important for the TTS system is the statistical probability of a morphological unit to productively contribute to word formation, and the system should be equipped to model this process.

The second important element of the previously quoted definition of productivity is that productive processes can in principle generate an unlimited number of new words. This observation is matched by the characterization of the unknown word analysis

noun forming prefixes					noun forming suffixes				
	N	Ftyp	n1	P		N	Ftyp	n1	P
*schwind-	1	1	1	1	-chen	1140	255	42	0.0368
vor-	104	14	2	0.0192	-ling	278	20	3	0.0108
be-	600	6	1	0.0017	-heit	604	7	2	0.0033
ge-	8125	164	10	0.0012	-schaft	11109	171	15	0.0014
semi-	12	3	0	0.0000	-ett	51	1	0	0.0000
adjective forming prefixes					adjective forming suffixes				
	N	Ftyp	n1	P		N	Ftyp	n1	P
*wiss-	1	1	1	1	-haft	1107	102	14	0.0126
ur-	108	10	1	0.0093	-voll	132	6	1	0.0076
un-	10010	601	64	0.0064	-är	502	17	1	0.0020
in-	219	49	1	0.0046	-lich	32168	569	51	0.0016
aller-	42	2	0	0.0000	-ig	3966	40	3	0.0008
verb forming prefixes					verb forming suffixes				
	N	Ftyp	n1	P		N	Ftyp	n1	P
weit-	94	11	3	0.0318	-er	65	24	5	0.0769
vor-	1401	31	4	0.0029	-el	1197	86	11	0.0092
ent-	13007	200	18	0.0014	-isier	1019	75	7	0.0069
ver-	53899	930	71	0.0013					
dar-	1071	6	1	0.0009					

Table 1: Selected noun, adjective, and verb forming affixes, their token (N) and type (Ftyp) counts, the number of hapax legomena (n1), and the computed Baayen productivity index (P). Higher values of P ($0 \leq P \leq 1$) indicate greater productivity. Affixes marked by an asterix (*) reflect an artifact due to the computation of P; they are in fact unproductive.

module as over-generating (see Discussion).

A statistical measure of productivity, which helps differentiate between degrees of productivity, has been proposed by Baayen [1]. His approach is based on the observation that the proportion of hapax legomena is much higher for intuitively productive affixes than for unproductive ones. Hapax legomena are usually defined as morphemes that occur in only one lexical expression in a given language; examples are English *cran* in *cranberry*, or German *brom* in *Brombeere*. In the context of this study, we define a hapax legomenon relative to a text database instead of to a lexicon. Given a particular morpheme, we list all word types, and their frequencies, in the database that are formed by this morpheme; a hapax legomenon is a—morphologically complex—word type with a token count of 1.

Following Baayen's suggestions, the productivity index (P) of a morpheme can be expressed as the ratio of hapax legomena (n1) to the total number of tokens containing that morpheme in the database (N) (see [8, pp. 4–9] for a more detailed discussion): $P = n1/N$.

Baayen and Lieber [2] applied this productivity measure to the English CELEX database [4]. The productivity estimates obtained in our study (see Section 2.2.) were derived from the German CELEX database, a fact that makes the criteria and thresholds quite comparable to those used by Baayen and Lieber. German CELEX is based on a text corpus of 6 million words. It contains 51,000 lemmata and 350,000 inflected forms, all of which are annotated for syntactic, morphological, and phonological properties. Raw frequency data are also available for both lemmata and word forms on the basis of the text corpus.

2.2. Productive Affixes

For the purposes of the present study, all prefixes and suffixes were extracted from the German CELEX database that occur as components in morphologically complex noun, adjective, and verb lemmata. For each affix the following data were computed: (1) token count for lemmata formed by the affix; (2) type count for lemmata formed by the affix; (3) number of types with 1 token (hapax legomena); (4) Baayen's productivity index. To get a more comprehensive picture of word formation processes in German, several pieces of linguistic information were manually added to each productive or marginally productive affix:

- stem type: whether or not the affix has any restrictions as to the type of stem it attaches to (e.g., latinate stems: *gener+ator, honor+ar*)
- umlaut: whether or not the affix causes the stem vowel to be umlauted (e.g., *Kunst* → *Künst+ler*)
- allomorphy: whether or not the affix causes some form of allomorphic variation of the stem
- infix: whether or not the affix (suffixes only) triggers infixation in compounds
- semantic function of the affix: negation, action, diminutive, location, and others

This additional linguistic information is exploited in various text analysis subcomponents of the TTS system. For instance, allomorphic variations are modeled in the morphological paradigm for each lexical entry. Attachment restrictions could be used in a more elaborate morphological word model (see Discussion).

Table 1 lists selected noun, adjective, and verb forming affixes, their token and type counts, the number of hapax legomena, and the computed productivity index. Higher values of P ($0 \leq P \leq 1$) indicate greater productivity. An index of 0 indicates that the affix is unproductive. Note that several affixes yield P = 1 because their token *and* type *and* hapax counts are 1, thus reflecting an undesired artifact inherent in the computation of P. Further inspection revealed that these cases typically involve morphemes which, although they are labeled as affixes in the database, should really be characterized as stems or roots. Examples are *schwind-* in the lexical entry *Schwindsucht* “consumption” or *wiss-* in *wissbegierig*. In general, affixes with a token count of less than 100 were ignored in the productivity study. However, a number of affixes which, on statistical grounds, were found to be unproductive were nevertheless included in the list of affixes in the declarative grammar of the morphological structure of German words (see Section 3.). Native speaker intuition overruled statistical evidence in these cases. It is also important to realize that the productivity index depends heavily upon the particular database it is applied to, and upon the degree of representativeness of that database relative to a (hypothetical) comprehensive written or spoken language corpus.

3. WORD MODEL

Our compositional model of German words is based on the morphological structure of words and the phonological structure of syllables. It has been implemented as a finite-state transducer using Richard Sproat's *Lextools* [15], a toolkit for compiling finite-state machines from linguistic descriptions. The module is therefore compatible with the other text analysis components in the German TTS system, all of which were developed in the same FST technology framework. One of the *Lextools,* the program *arclist*, is particularly well suited for the morphological analysis of compounds and unknown words. The tool facilitates writing a finite-state grammar that describes words of arbitrary morphological complexity and length.

The core of the module is a list of approximately 5000 nominal, verbal, and adjectival stems that were extracted from the morphologically annotated lexicon files of the TTS system. To this collection we added about 250 prefixes and 220 suffixes that were found to be productive or marginally productive in the previously described study. We also included 8 infixes (*Fugen*) which German word formation grammar requires as insertions between components within a compounded word in certain cases, such as *Arbeit+s+amt* “employment agency” or *Sonne+n+schein* “sunshine”.

The two most important aspects of this linguistic description are, first, the decision which states can be reached from any given current state and, second, which of the legal paths through the graph should be preferred over other legal paths. The first aspect can be regarded as an instantiation of a declarative grammar of the morphological structure of German words. The second aspect reflects the degrees of productivity of word formation, represented by costs on the transitions, or *arcs*, between states.

Figure 1 shows segments of the *arclist* source file for unknown word decomposition. In each line, the first and second column are labels for two states, the state of origin and the state of destination, respectively. The third column contains a string of symbols on the arc between the two states. These strings consist of regular orthography, annotated with lexical and morphological labels including morpheme boundaries (+), symbols for primary lexical stress ('), and an optional cost for the transition.

```
START   PREFIX  [Eps]
START   PREFIX  er [pref] [+] (0.3)
START   PREFIX  ge [pref] [+] (0.3)
START   PREFIX  'un [pref] [+] (0.3)
...     ...     ...
PREFIX  START   [Eps] (1.0)
PREFIX  ROOT    SyllableModel [root] (10.0)
PREFIX  ROOT    d'ampf [root]
PREFIX  ROOT    f'ahrt [root]
PREFIX  ROOT    f'ind [root]
PREFIX  ROOT    kapit'än [root]
PREFIX  ROOT    st'ell [root]
...     ...     ...
ROOT    PREFIX  [comp] [+] (0.1)
ROOT    SUFFIX  [Eps] (0.2)
ROOT    SUFFIX  [+] lich [suff] (0.2)
ROOT    SUFFIX  [+] keit [suff] (0.2)
ROOT    END     [noun] (1.0)
...     ...     ...
SUFFIX  ROOT    [Eps] (0.2)
SUFFIX  END     [noun] (1.0)
SUFFIX  FUGE    [+] s [fuge] (0.2)
SUFFIX  FUGE    [+] n [fuge] (0.2)
...     ...     ...
FUGE    START   [comp] [+] (0.5)
END
```

Figure 1: Segments of a declarative grammar (in *arclist* format) for unknown word decomposition in German. Column 1: state of origin; column 2: state of destination; column 3: string on arc between states with optional cost.

The transition from the initial state START to the state PREFIX is defined by a family of arcs that represent productive prefixes. One of the prefix arcs is labeled with [Eps] (Epsilon, the empty string), allowing for words or components of complex words that do not have a prefix. Multiple prefixes can be modeled by returning from PREFIX to START; by assigning a relatively high cost to this path, analyses which require only one prefix are favored. The transition back to START carries the [Eps] label.

A large family of arcs from PREFIX to ROOT represents nominal, verbal, and adjectival stems. Sequences of stems are modeled by returning to PREFIX from ROOT without intervening affixes; in this case, the tag [comp] indicates the end of a complete subcomponent in a complex word. If a word terminates in a stem, the appropriate path through the graph is from ROOT to END.

The transition from PREFIX to ROOT that is labeled *Syllable-Model* is a place holder for a phonetic syllable model which reflects the phonotactics and the segmental structure of syllables in German, or rather their correlates on the orthographic surface.

This allows the module to analyze substrings of words that are unaccounted for by the explicitly listed stems and affixes in arbitrary locations in a morphologically complex word. Applying the syllable model is expensive because we want to cover the orthographic string with as many known components as possible. The costs actually vary depending upon the number of syllables in the residual string and the number of graphemes in each syllable. For the sake of simplicity we assign a flat cost of 10.0 in our example. The syllable model is described in more detail in Section 4.

Productive suffixes are represented by a family of arcs between the states ROOT and SUFFIX. Analogous to the case of prefixes, one such arc carries the empty string to skip suffixes, and consecutive suffixes are modeled by returning to ROOT from SUFFIX. The iteration over suffixes is less expensive than the one over prefixes because sequences of suffixes are significantly more common than multiple prefixes. If the word ends in a suffix, the path through the graph continues from SUFFIX to END.

The last family of arcs represents the *Fugen* infixes as transitions from SUFFIX to FUGE. There is only one legal continuation through the graph from FUGE, viz. back to the beginning of the graph. This design reflects the fact that *Fugen*, by definition, can only occur between major subcomponents of a complex word, each of which have their own stem; hence also the indication of the completion of the subcomponent by means of the tag [comp].

On termination the machine labels the word as a noun by default. In a more sophisticated word model it might be possible to assign the part-of-speech category of the unknown word on the basis of the types of stems and affixes involved, by distinguishing between noun, verb, and adjective forming affixes. However, as of now we are lacking the capability to disambiguate concurrent analyses, which are very likely to occur because many stems and affixes are ambiguous in terms of their part-of-speech status. In the current implementation it is sufficient for the prosodic components of the TTS system to know that the word is a content word, which is a safe assumption for novel words.

Most arc labels are weighted by being assigned a cost. Weights are a convenient way to describe and predict linguistic alternations. In general, such a description can be based either on an expert's analysis of linguistic data and his or her intuition, or on statistical probabilities derived from annotated corpora. Works by Riley [11] and Yarowsky [16] are examples of inferring models of linguistic alternation from large corpora. However, these methods require a database that is annotated for all relevant factors. Despite our large *raw* corpus, we are lacking the type of database resources required by these methods. The weights in the unknown word analysis module of our German TTS system are currently based on our analysis of productive word formation processes presented above or on linguistic intuition. Weights are assigned such that direct matches of input strings to entries in the lexicon will be less expensive than unknown word analysis. There is no legal path through the unknown word graph that comes at zero cost; the minimal cost of 1.0 would be for a simplex stem that is explicitly listed and does not have any affixes.

The morphological information provided by the unknown word analysis component is subsequently used by the phonological, or pronunciation, rules.

class	description	phones
P	unvoiced stops	ʔ p t k
B	voiced stops	b d g
S	unvoiced fricatives	f s ʃ ç x h
Z	voiced fricatives	v z ʒ
N	nasals	m n ŋ
L	liquids	l r R
G	glides	j
V	vowels, diphthongs	i: y: e: ø: ɛ: u: o: a: aɪ aʊ ɔʏ ɪ ʏ ɛ œ ʊ ɔ a ə ɐ

Table 2: German phone classes.

4. SYLLABLE MODEL

In general, German pronunciation is more sensitive to morpheme boundaries than to syllable boundaries. This observation is reflected in the implementation of the phonological component of our TTS system. The TTS lexicon is extensively annotated with morphological information, which is used in the context specifications of pronunciation rules. Morpheme boundaries in German tend to also be syllable boundaries (but not the other way around!), with the general exception of inflectional affixes. Note that other components of the TTS system, e.g. the segmental duration model, do rely on syllable boundary information. The main syllabification algorithm operates on the output of the phonological component, i.e., on the sequence of phonemes and syllabic stress symbols (see [6] for details). A variant of this syllabifier, operating on the lexically and morphologically annotated orthographic surface, has been integrated into the word model; the design considerations are described in this Section.

4.1. Phonotactics

The phonotactics of German allow complex consonant clusters in both the onset and the coda of syllables. The maximum number of consonants in the onset is 3, in the coda 5 (or 6, if contractions like *du schrumpfst's* [du: ʃrʊmpfsts] "you shrink it" are considered, too). However, certain restrictions exist as to which consonant or class of consonants can occur in which position within the onset or coda. For instance, there are only four possible onset clusters with three consonants: SPL, PSL, PSZ, and SPZ (see Table 2 for an explanation of the symbols); and no phones other than obstruents can occur *before* an obstruent in the onset. In the coda, only combinations and alternations of P and S can occur in positions 3 through 5 (or 6); also, after the first obstruent, no phones other than obstruents can occur in the coda.

Sonorants (nasals, liquids, and glides) can only occur adjacent to the syllable nucleus. This pattern is sometimes referred to as the sonority principle, which ranks phone classes according to their natural acoustic sonority, which in turn is a correlate of the degree of constriction of the vocal tract. A typical nominal scale from most to least sonorant is: open vowels, closed vowels, glides, liquids, nasals, voiced fricatives, unvoiced fricatives, voiced stops, unvoiced stops.

Arguably, this tentative ranking follows more the linguist's intuition and his or her knowledge of the syllable structure in many languages with consonant clusters, rather than actual acoustic measurements. It is therefore important to keep in mind that it is impossible to give a language independent sonority ranking that is based on both acoustic properties and phonotactics. For instance, while *alm, arm, art* are legal syllables in German and **aml, *amr, *atr* are not, this situation may be different, or even reversed, in other languages. In fact, *amr* [ʔamr] "command" and *naml* [naml] "ant", are legal syllables in Arabic, and the name *Pjotr* [pjɔtr] is analyzed and perceived as monosyllabic in Russian.

The complexity of syllable onset and coda structure in German poses serious problems for a syllabification algorithm because, despite of the above-mentioned restrictions, we more often than not observe a high degree of ambiguity and multiple alternative syllable boundary locations in polysyllabic words, notably in compounds. Determining the syllable boundary is important in German TTS because the pronunciation of most phonemes is a function of their position in the syllable relative to the syllable boundaries. This is most evident in the case of phonologically voiced obstruents: the voicing opposition for stops and fricatives is neutralized in the syllable coda. The phonological minimal pair *Bund* "union" – *bunt* "colorful" is in fact homophonic: [bʊnt].

The task of the syllable model is to parse orthographic substrings *within* morphologically complex unseen words, typically compounds and names, substrings that cannot be decomposed into the explicitly listed morphemes of the word model. Under the assumption that all productive or marginally productive morphemes of German are covered by the explicit list, unknown substrings are very likely to be *lexical* morphemes that happen not to be accounted for. Therefore, it is relatively safe to mark with morpheme boundaries the transitions from a known morpheme to the unknown substring and from that substring to another known morpheme. Similarly, any syllable boundary assigned by the syllable model *within* an unknown substring is treated as a morphological boundary as a first approximation. It should be pointed out, however, that the latter assumption is somewhat risky; I will come back to this issue in the Discussion.

4.2. Syllable Structure

The description of the phonetic manifestation of German syllable structure is rooted in both theory and empirical data. The theory can be found in text books on German phonetics and phonotactics (e.g., [7]). The empirical data were derived from the German Celex database [4]. We extracted transcriptions of all monosyllabic words from the word forms database, converted each phone symbol into the symbol of the phone class it belongs to (see Table 2) and obtained a list of onset and coda types sorted by frequency. This procedure yielded 534 unique syllable types, 25 onset types, and 54 coda types. Note that the number 534 refers to actually observed syllable types, which is about 40% of the phonotactically possible syllable types (25 * 54 = 1350).

The unknown word model transducer, as a text analysis component in the TTS system, takes orthographic input. Two transformations had to be performed to implement the phone class based syllable structure as an *arclist*-type source file for a finite-state machine. First, every phone class was expanded to the set of phones that are members of that class (see Table 2). Second, the phone symbols were replaced with all the graphemes and grapheme strings that have a possible pronunciation corresponding to the phone in question. This operation can be characterized as the inversion of applying pronunciation rules. In fact, since the pronunciation rules in our system are implemented as finite-state transducers, it would have been possible to invert the direction of the transduction. However, due in part to the large number and high complexity of context specifications in the rules, and in part to the fact that the task was not to find the one best (and correct) mapping from one representation to another but instead to *all* possible correspondences, it turned out to be more practical to simply enumerate the orthographic correlates of each phoneme. By exploiting the lexical database, the procedure became largely automatic. The resulting transducer was incorporated into the unknown word analysis module (*SyllableModel* in Figure 1).

5. APPLICATION TO NAMES

The approach to unknown word decomposition described above has also been applied to the analysis of names in our German TTS system. Arguably, names are not equally amenable to morphological processes, such as word formation and derivation, or to morphological decomposition, as regular words are. That does not render such an approach unfeasible, though, as was shown in a recent evaluation of the system's performance on street names. Street names are an interesting category because they encompass aspects of geographical and personal names. In our study [5], we reported a pronunciation error rate *by word* of 11–13% for unknown names. In other words, roughly one out of eight names is pronounced incorrectly.

This performance compares rather favorably with results reported in the literature, for instance from the German branch of the European Onomastica project [10]. Onomastica was funded by the European Community from 1993 to 1995 and aimed to produce pronunciation dictionaries of proper names and place names in eleven languages. The final report describes the performance of grapheme-to-phoneme rule sets developed for each language. For German, the accuracy rate by word for quality band III—names that were transcribed by rule only—was 71%, yielding an error rate of 29%. The grapheme-to-phoneme conversion rules in Onomastica were written by experts, based on tens of thousands of the most frequent names that were manually transcribed by an expert phonetician. In our TTS system, the phonological or pronunciation rules capitalize on the extensive morphological information provided by annotated lexica *and* the unknown word analysis component.

One obvious area for improvement is to add a name-specific set of pronunciation rules to the general-purpose one. Using this approach, Belhoula [3] reports error rates of 4.3% for German place names and 10% for last names. These results are obtained in recall tests on a manually transcribed training corpus. The addition of name-specific rules presupposes that the system knows which orthographic strings are names and which are regular words. The problem of name detection in arbitrary text (see [14] for an approach to German name tagging) has not been addressed for our TTS system so far. It is by-passed for the time being by integrating the name component into the general text analysis system and by adjusting the weights appropriately.

6. DISCUSSION AND FUTURE WORK

We presented the implementation of an unknown word analysis module as a central component of text analysis in the Bell Labs German text-to-speech system. The module integrates a model of the morphological structure of words and the productivity of word formation with a model of German syllable structure. It has been shown to be capable of correctly analyzing morphologically complex words, such as (novel) compounds and names. The analysis produces a rich morphological annotation for unseen words that is comparable to the annotation of words that are actually found in the lexicon of the TTS system. This is important for the applicability of general-purpose pronunciation rules in the phonological component of the system.

Several issues remain to be investigated more closely. Most importantly, the FST machinery is over-generating and tends to be over-analyzing, occasionally with undesired side effects on pronunciation rules. One way to prevent over-generation is to implement an even more refined morphological word model that applies phonological, syntactic, or semantic restrictions as to which affixes attach to which (type of) stems, and which sequences of affixes without intervening stems are permitted. This approach would require studies on a sizable annotated text corpus.

Undesired side effects on pronunciation arise from the fact that currently any syllable boundary assigned by the syllable model *within* an unknown substring is treated as a morphological boundary. This design is a best-guess approximation, and it is somewhat risky because morpheme boundaries trigger certain phonological rules. For example, the street name *Rimparstraße* is incorrectly analyzed as *Rim+par+straße*, with the pronunciation computed as [riːmpaɐʃtraːsə] by applying a phonological rule stating that the grapheme <i> is pronounced [iː] when followed by exactly one consonant and a morpheme boundary. As it were, *Rimpar* happens to be a proper name that cannot be further decomposed. The correct pronunciation is [rɪmpaɐʃtraːsə] because <i> is pronounced [ɪ] when followed by two or more consonants without an intervening morpheme boundary; the syllable boundary, assumed to occur between <m> and <p>, does not trigger the same change in vowel quality and quantity as a morpheme boundary does. The example shows that the optimal depth of analysis has not yet been determined, and it is somewhat unclear whether and how it can be found at all.

Another issue that deserves further study is the assignment of syllabic stress to morphologically complex unknown words. Currently, primary lexical stress is assigned to the first stem of the word as a rule. However, certain affixes have a strong tendency to attract primary stress; examples are the prefix *herab-* as in *herabfallen* or the suffix *-tion* as in *Produktion*. If we see this class of affixes, it is relatively safe to assume that the primary stress falls on the affix and not on the stem. The overwhelming majority of cases is, of course, less clear-cut, and we observe quite a number of words with components competing for primary stress. Furthermore, in longer compounds it is often desirable to assign primary lexical stress to more than one syllable. For the time being, and without a thorough corpus-based study, the assignment of only one lexical stress per morphologically complex word is the most reasonable approach.

7. REFERENCES

1. Baayen, H. On frequency, transparency and productivity. *Yearbook of Morphology 1992* 1993, 181–208.
2. Baayen, H., and Lieber, R. Productivity and English derivation: a corpus based study. *Linguistics 29* 1991, 801–843.
3. Belhoula, K. A concept for the synthesis of names. In *ESCA Workshop on Applications of Speech Technology* (Lautrach, Germany, 1993).
4. Celex. The CELEX lexical database—Dutch, English, German. CD-ROM, 1993. Centre for Lexical Information, Max Planck Institute for Psycholinguistics, Nijmegen.
5. Jannedy, S., and Möbius, B. Name pronunciation in German text-to-speech synthesis. In *Proceedings of the Fifth Conference on Applied Natural Language Processing (ANLP-97)* (Washington, DC, 1997), Association for Computational Linguistics (ACL), pp. 49–56.
6. Kiraz, G. A., and Möbius, B. Multilingual syllabification using weighted finite-state transducers. In *Proceedings of the Third ESCA Workshop on Speech Synthesis* (Jenolan Caves, Australia, 1998).
7. Kohler, K. J. *Einführung in die Phonetik des Deutschen*, 2nd ed. Erich Schmidt Verlag, Berlin, 1995.
8. Lieber, R. *Deconstructing morphology: Word formation in syntactic theory*. University of Chicago Press, 1992.
9. Möbius, B., Sproat, R., van Santen, J., and Olive, J. The Bell Labs German text-to-speech system: an overview. In *Proceedings of the European Conference on Speech Communication and Technology (Eurospeech)* (Rhodos, Greece, 1997), vol. 5, ESCA, pp. 2443–2446.
10. Onomastica. Multi-language pronunciation dictionary of proper names and place names. Tech. rep., European Community, Linguistic Research and Engineering Programme, European Commission, DG XIII, 1995. Project No. LRE-61004, Final Report, 30 May 1995.
11. Riley, M. D. Tree-based models of speech and language. In *Proceedings of the IEEE-IMS Workshop on Information Theory and Statistics* (Alexandria, VA, 1994).
12. Schultink, H. Produktiviteit als morfologisch fenomeen. *Forum der Letteren 2* 1961, 110–125.
13. Sproat, R., Ed. *Multilingual Text-to-Speech Synthesis: The Bell Labs Approach*. Kluwer Academic, Dordrecht; Boston, MA; London, 1998.
14. Thielen, C. An approach to proper name tagging in German. In *From text to tags—Issues in multilingual language analysis. Proceedings of the ACL SIGDAT Workshop* (University College, Belfield, Dublin, Ireland, 1995), pp. 35–40.
15. van Santen, J., and Sproat, R. Methods and tools. In Sproat [13], pp. 7–30.
16. Yarowsky, D. Homograph disambiguation in speech synthesis. In *Proceedings of the Second ESCA/IEEE Workshop on Speech Synthesis* (New Paltz, NY, 1994), ESCA/IEEE, pp. 244–247.