Place assimilation across words in running speech: Corpus analysis and perception

Frank Zimmerer
Goethe Universität, Institut für Phonetik, Georg-Voigt–Str. 6/2, 60325 Frankfurt, Germany and Universität Konstanz, FB Sprachwissenschaft, D186, 78457 Konstanz, Germany

Henning Reetz
Goethe Universität, Institut für Phonetik, Georg-Voigt–Str. 6/2, 60325 Frankfurt, Germany

Aditi Lahiri[a)]
Universität Konstanz and Centre for Linguistics and Philology, University of Oxford, Walton Street, Oxford OX1 2HG, United Kingdom

(Received 18 July 2007; revised 2 August 2008; accepted 20 October 2008)

Place assimilation can lead to neutralization of segmental contrasts. It is controversial, however, to what extent such neutralizations actually happen in natural speech. This study examines: (i) the degree to which regressive place assimilations occur in word final consonants in conversational German, and (ii) whether these assimilations are perceived as neutralized by listeners. The production analysis, based on spontaneous speech, shows that complete assimilations do take place in conversational speech and that there is a clear asymmetry between coronal versus labial and dorsal segments. Furthermore, function words show a higher degree of assimilation than lexical words. Two experiments examined the effects of assimilation on perception. A forced choice reaction time perception experiment, using nasal stimuli from the corpus, examined how fast and accurately listeners identified sounds in different segmental contexts. Results indicate that (a) with equal accuracy and speed, listeners identified original and assimilated [m]s; (b) unassimilated-/m/s were identified equally well across contexts, but not unassimilated-/n/s. A free transcription experiment reproduced these findings. An acoustic analysis provides further evidence that regressive place assimilation across word boundaries can result in absolute neutralization of place contrasts in running speech. The results support models predicting asymmetries between coronal versus labial and dorsal consonants. *© 2009 Acoustical Society of America.* [DOI: 10.1121/1.3021438]

PACS number(s): 43.70.Mn, 43.71.Es, 43.71.Sy [PEI] Pages: 2307–2322

I. INTRODUCTION

Speech is variable—certainly across speakers and across dialects, but also within a single speaker depending on speech rate or style. Acoustic shapes of lexical items easily vary across contexts or sentence types. Within a model of discrete phonological entities, it is customary to assume that reductions, deletions, insertions, and assimilations are processes which modify or alter idealized lexical representation in running speech. A German word like *einverstanden* "agree-past participle" occurring 47 times in the Kiel corpus (IPDS, 1994) has 23 different variants in the database (for a complete list see Appendix A).[1]

There is no utterance in the database which exactly matches the canonical pronunciation (i.e., [ˈʔaɪnfɐˌʃtandən]), most cases exhibiting more than one deviation. Not only are there many types of variation, but the deviations from the norm are optional and need not be complete and may still be perceptible. Remnants of a deleted sound may still be present as in [ˈa̰ɪ̰nfɐˌʃtan̰n], [ˈa̰ɪ̰nfɐˌʃtanʔn], or [ˈa̰ɪ̰nfɐˌʃtan̰]), where some segments have been deleted completely, or glottalization indicates that a stop (i.e., [d]) has been severely reduced.[2] Transcriptions of place assimilations suggest complete neutralization of a featural contrast, as in [ˈaɪmfɐˌʃta̰nn] where additionally to deletions, the [n] is assimilated to the labiality of [f].[3]

However, as in deletions, traces of the original [n] may still be found in the signal. For instance, Nolan (1992) argued that assimilations were more likely to be gradient than complete, and that target information was available in assimilated sequences (see also Gow, 2002; for voicing assimilation, see Snoeren *et al.*, 2006). Listeners are sensitive to these gradient assimilations in production and could identify residual alveolar gestures in 40% of the assimilated tokens (Nolan, 1992, p. 271). Gow's results also indicate that listeners use the information of the underlying place of articulation even in segments that auditorily sound as if they are completely assimilated (Gow, 2002; Dilley and Pitt, 2007; also Manuel, 1995; Snoeren *et al.*, 2006). Indeed, some researchers express doubt concerning the very existence of complete assimilation (Gow, 2002; see also Snoeren *et al.*, 2006).

However, in a recent extensive coverage of regressive assimilation of naturally spoken American English (Buckeye Corpus of Conversational Speech, Pitt *et al.*, 2006), Dilley and Pitt (2007) found that 9% of coronal (alveolar) word final stops and nasals were transcribed as assimilated to the place of articulation of the following consonant (labials and

[a)]Electronic mail: aditi.lahiri@ling-phil.ox.ac.uk

velars).[4] Acoustic measurements consisting of the change in F2 and amplitude of the preceding vowel showed that these frequently did not differ between the assimilated consonants and the canonical labials and velars. They conclude that "assimilation is often complete or nearly complete in spontaneous speech" (Dilley and Pitt, 2007, p. 2350). One must note, however, that the F2 values were gradient for both consonants labeled as assimilated, as well as those in an assimilatory context (i.e., followed by labials or velars) as compared to alveolars in a nonassimilatory context (i.e., followed by other alveolars). As the authors report, a possibility exists that the real number of assimilations is underestimated, since even some instances of those that were labeled as unassimilated could be actually assimilated because the labelers are always reasonably conservative.

Thus, in spite of the optional and gradient nature of fast speech processes, neutralization due to assimilation can be perceived as complete. Furthermore, such assimilations can even lead to orthographic changes. Orthography tends to be conservative and even if a pronunciation change has occurred, the spelling remains often unaltered. However, when the orthography changes (without formal institutional intervention such as the German *Rechtschreibreform*), one is reasonably sure that a change has really taken place. For instance, words with the negative prefix {in-} have been borrowed into English from Romance at different times. A word like *impossible* could be spelt earlier as ⟨inpossible⟩: *It es bot foli al pi talking, And als an inpossibile thing* **1300** *Cursor M.* 14761 (OED, 1989, p. 732). The ⟨n⟩ is now always pronounced as a labial and this place assimilation changing the [n] of {in-} to [m] when labial [p, b, m] follow is now also always reflected in spelling.[5] Listeners must have perceived the assimilation to lead to a change in the (conservative) orthography. A new formation like *input*, which is not made with the negative prefix, preserves the ⟨n⟩ in the spelling, although it is also pronounced with a [m].

Although there is little doubt that dictionary-like pronunciations are not the norm in connected speech and there is an increasing number of spoken language corpora which are used in publications, such as Snoeren and colleagues (2006) and Dilley and Pitt (2007), there is still a dearth of statistically reliable data as to what extent connected speech phenomena like assimilations actually occur in other languages and moreover, even less is known about how they are perceived by normal listeners and trained labelers. In this paper, we analyze normal running speech to quantify how often the contexts for assimilations arise, how frequently such assimilations are realized, and how they are perceived, focusing on German where there is no study as yet bringing together corpus analysis and the repercussions for perception. In the following sections we first discuss possible across-word assimilation phenomena in German and provide a complete analysis of such assimilations in the Kiel corpus. Our analysis is first divided into a section for function words and a separate section for lexical words. There is ample evidence that the phonological and phonetic behavior of these two word categories is different (e.g., Selkirk, 1984; Kaisse, 1985; Nespor and Vogel, 1986; Hall, 1999; Ogden, 1999; Philipps, 2001; Local, 2003; Kabak and Schiering, 2006; Bybee, 2007). For instance, function words in German are often drastically reduced (Hall, 1999). For English, it has been reported that /m/ in the function words "I'm" may assimilate to neighboring segments, whereas /m/ in content words such as "time" does not (Ogden, 1999; Local, 2003). We also examine whether function words behave differently when it comes to regressive place assimilation and compare the findings of the two sections. This analysis is followed by two perception experiments examining the ability of listeners to hear the altered sounds based on the Kiel transcriptions. We offer a proposal for modeling alterations by assimilations in a framework assuming underspecification, such as the Featurally Underspecified Lexicon model (FUL) (Lahiri and Reetz, 2002).

II. ASSIMILATIONS IN GERMAN

How inviolable are word final consonants? Such consonants are of course subject to change, either by progressive assimilation within a word or by regressive assimilation across words. Although progressive place assimilations are reported to be frequent within a word in German, cf. *geben* [gebən]>[gebm], regressive assimilation across words is more controversial (Wurzel, 1970; Dressler *et al.*, 1972; Benware, 1986; Hall, 1992; Wiese, 1996; but see Kohler, 1995). As Wiese (1996) states, when it is possible to pronounce two words as a single unit then regressive assimilation is more likely (cf. *man kommt* "one comes" pronounced as [maŋ kɔmt] versus no assimilation in *der Mann kommt* "the man comes" [man kɔmt]). More definite conclusions regarding regressive assimilations in German is difficult since in his words, "…first, there is little systematic study of such differences, and, second, at the tempo of fast speech, assimilation is certainly possible in the latter example" (Wiese, 1996, p. 221). Nonetheless, regressive assimilations across words are not unknown and the possibility is at least mentioned by most of these authors.[6]

Kohler, however, explicitly claims that regressive place assimilation takes place across word boundaries (Kohler, 1995, p. 206; see also Kohler, 1990) and cites several examples where such assimilations occur. One such example is *bunt machen* "to make colorful" [bʊnt maxŋ] being pronounced as [bʊmp maxŋ]. A study on the Viennese variety of German by Dressler and his colleagues also mentions the possibility of regressive place assimilation in fast speech (Dressler *et al.*, 1972).

Given that there is little systematic work on regressive assimilation across word boundaries in conversational German, we turned to the Kiel corpus for natural speech data. The Kiel corpus (IPDS, 1994) provides us with ideal data for examining minute phonetic variations in naturally spoken dialogues. The corpus makes available detailed phonetic transcriptions and hand labeled segmentation of the acoustic signal performed by trained phoneticians. Thus, we have rich material to study across-word variations as noted by phoneticians who used both the information in the signal as well as their own perception to make decisions concerning what was actually produced.

In what follows, we refer to the sequence of consonants across word boundaries as C_1 and C_2. The word final segment (C_1), which could assimilate, will be referred to as the target and the word initial segment of the following word (C_2) as the trigger as shown in (1). C_1 could be any stop, fricative or nasal in German, whereas C_2 could be any obstruent or nasal which may occur in that position. Word finally (target position), voiced stops and fricatives are devoiced regularly in German (*Auslautverhärtung*—final devoicing—see Kohler, 1995; Wiese, 1996; Hall, 2000, and references therein). Consequently, in production, there are no word-final voiced obstruents.[7]

(1) TARGET (C_1) and TRIGGER (C_2) in word sequences
$[\ldots C_1]_{\omega 1}[C_2 \ldots]_{\omega 2}$

e.g., *Termin* *bleiben* ("appointment stick" —stick to appointment)

The following issues are addressed in the analyses of the speech data: (a) How often do German speakers produce regressive assimilations across words? (b) Is there a particular place of articulation for C_1 which favors assimilation? For instance, are [coronal] sounds more likely to assimilate than [labial] ones? (c) Does the manner of articulation of C_1 matter for regressive assimilation? For example, do nasals assimilate more often than stops in running speech? (d) Does the place and manner of articulation of the C_2 trigger correlate with regressive assimilation? (e) Does the lexical status of the first word (function words versus lexical words) increase the probability of assimilation since function words are supposed to be less stable and more vulnerable to alterations?

Other than quantifying the number of assimilations in conversational speech, we also address the issue of completeness of these assimilations in perception. Two perception studies were conducted using selected material from the database. The first experiment used a forced choice phoneme identification task on fragments of words from selected dialogues. In a second experiment, subjects were asked to transcribe freely what they heard. Our goal was to observe how listeners would perceive segments labeled as assimilated in the speech corpus and whether there were any remnants of the original segment to affect the speed and accuracy of identification as compared to unchanged segments.

III. ANALYSIS OF SPONTANEOUS SPEECH

The Kiel corpus consists of about 4 h of dialogues of 42 (northern) German speakers (18 female, 24 male) who were engaged in an appointment-making task. They were each given a schedule with existing and sometimes conflicting appointments, and their task was to decide on future meetings. In order to ensure a high degree of natural speech, the speakers were ignorant of the schedule of their partners. The dialogues were recorded with the speakers placed in different sound-treated rooms communicating by headsets. As mentioned earlier, all dialogues were transcribed and labeled by trained phoneticians using visual scaleable spectrograms and oscillogram displays as well as auditory information (Kohler *et al.*, 1995, p. 33). Other than the phonetic transcription of what was actually pronounced, the corpus contains an orthographic text, along with its canonical phonetic transcription. The idealized canonical transcription denotes how utterances should be if they were spoken in accordance with a careful dictionary-like pronunciation. This allows a direct comparison of a canonical transcription with the actual—phonetically transcribed—pronunciation. The nature of the task and the fact that participants had to make very similar appointments restricted the vocabulary. For instance, dates, times, and days of the week occur very often. Nevertheless, since the speakers were unaware of the purpose of the recordings, the conversations were very natural.

A. Material and methods

Overall the *Kiel Corpus of Spontaneous Speech* consists of 1984 turns of dialogues by 42 speakers available in three CDs. All transcribed dialogs from the appointment-making task consisting minimally of two words, including the test dialogs, went into our analysis. The longest turn is about 54 s and the shortest ones are less than a second. An example of average length of one turn is given in (2) with (a) the orthographic transcription with our English translation, (b) the canonical transcription in IPA, and (c) the phonetic transcription as given in the Kiel corpus (omitting instances of nonspeech sounds such as clicking, smacking, etc.).

(2) Example of a dialogue in the Kiel corpus (utterance g071a006):

(a) Orthographic transcription *Das würde mir ganz ausgezeichnet passen. Machen wir das fest*? "That would be excellent for me. Can we confirm this?,"
(b) Canonical transcription (IPA transcription of the orthographic form) [das vʏɐdə mɪːɐ ˈgants ˈʔaʊsgəˌtsaɪçnət ˈpasən. ˈmaxən viːɐ das ˈfɛst?],
(c) IPA transcription of the actual pronunciation [das vʏɐd mɪːɐ ˈgants ˈãʊ̃sgəˌtsaɪhnət ˈpasn. ˈmaxm viːɐ das ˈfɛst?].

We counted all possible contexts of regressive place assimilations of nasals and obstruents, and then summed up all cases where they actually occurred, even across sentence boundaries. This meant that instances where the C_1-target and C_2-trigger had the same place of articulation were ignored. For the analysis, utterances were excluded where technical problems led to incomplete speech signals, or where speakers produced false starts, or where there were intervening hesitational markers as *ähm* or *m(hm)* ("ahem, hm") like in *machen äh(m) wir*. Furthermore, to rule out possible confounds, we did not include utterances where a possibility of progressive place assimilation existed and thus target and trigger could not be identified unambiguously. For example, the assimilated [m] in a phrase like *haben wir* ("have we") [haːbən viːɐ] spoken as [haːbm viːɐ] has two potential triggers, the preceding labial [b] or the following labial [v], and was therefore not considered in our data set. We also excluded words where the last segment (C_1) was

TABLE I. Obstruents and nasals in German and their phonological place features.

Labial	*bilabial, labiodental*	[m, p, b, f, v, pf]
Coronal	*alveolar, palatoalveolar, palatal*	[n, t, d, s, ʃ, z, ç, ts, tʃ]
Dorsal	*velar*	[ŋ, k, g, x]

deleted to rule out all possible confounds connected to deletions. Thus, phrases like *und Mittwoch* ("and Wednesday") [ʊnt ˈmɪtvɔx] pronounced without word final [t] as [ʊn ˈmɪtvɔx] were not included.[8] All obstruents and nasals were treated as possible triggers (C_2). The phonological features of the consonants that were taken into account, both as target and trigger, are given in Table I.[9]

The segments [ŋ, tʃ, s, x] do not occur word initially in standard German, except in a handful of loanwords. Due to final devoicing, one expects the voiced consonants to be devoiced word finally, but this does not affect place assimilation.

B. Analysis

Function words behave differently than content words in many different ways (e.g., Selkirk, 1984; Kaisse, 1985; Nespor and Vogel, 1986; Hall, 1999; Ogden, 1999; Philipps, 2001; Local, 2003; Kabak and Schiering, 2006; Bybee, 2007). Therefore, we split up the following analysis into a section for function words, another section for content words, and then we compare the results. The issue of interest to us is whether function words behave differently concerning regressive place assimilation. Since there is considerable controversy concerning which words count as function words, we opted for the classification in the Kiel corpus (marked with a final "+" in their transcription).

1. Function words

An overview of the different kinds of function words occurring in the database is given in (3). The function words could be either trigger or target. However, we ignored the syntactic category of the second word in a sequence, since our main point of interest is on the words that undergo assimilation, therefore the lexical status of the trigger was of no relevance in our analysis.

(3) Examples for different function word categories in the Kiel corpus:

(a) Auxiliaries: *bin, hatte, gewesen, möchte* ("am, had, been, would like"),
(b) Determiners: *der, die, das, ein, eine* ("the.*masc*, the.*fem*, the.*neut*, a.*masc*, a.*fem*"),
(c) Pronouns: *ich, wir, Sie, Ihre, Ihnen*, ("I, we, you.*hon*, you.*hon*.*gen*, you.*hon.acc*")
(d) Prepositions: *in, am, bis*, ("in, at.*dat*, until/to'"),
(e) Demonstratives: *diesen, dieser, diesem* ("this-*case*"),
(f) Conjunctions: *und, aber, zwar* ("and, but, but/namely").

Overall, 4144 function words qualified as target (C_1) in a sequence of two consonants at word boundaries. Out of those, in 266 (6.4%) instances the target C_1 was transcribed in the corpus as having been pronounced with a different place of articulation from the canonical form, e.g., *ein Montag* [aɪm moːntax] instead of [aɪn moːntax]. Table II(a)–(c) show the data for all occurrences of targets and the corresponding triggers, with the numbers and percentages of assimilated segments.[10]

The results clearly indicate that although regressive place assimilations are not obligatory, they do occur across word boundaries of function words. In Table II(b) we see that 232 out of 2961 [coronal] sounds assimilate in place to the following segment, most of which were /n/. Out of a total of 1230 /n/ final function words, 225 or 18.3%, were labeled as assimilated; 168 out of 1036 (16.2%) words ending in /n/ assimilated to [m] and 57 out of 194 (29.4%) changed to [ŋ], when followed by [labial] or [dorsal] consonants, respectively. Out of a total of 200 function words ending in /t/, only 4 assimilated to [p] when a labial followed, and none assimilated before [dorsal] segments. Overall, 1021 function words ended in /s/, one of which was assimilated to a labial [f]. Finally, out of 510 /ç/ final function words, 2 assimilated to [f].

Turning to the [labial] final function words [Table II(a)], there were in all 724 of which 27 assimilated, all of which were /m/. There were 583 instances of /m/ final function words and 27 were labeled as having changed its place of articulation, 23 (i.e., 4.3%) to [n] when followed by a [coronal], 4 (9.3%) to /ŋ/ when followed by a [dorsal]. None of the 82 /p/ or 59 /f/ final function words assimilated. As for the [dorsal] final function words [Table II(c)], they all ended in /x/, and 7 out of 459 instances (1.5%) showed assimilation—six times to [f] when a [labial] followed, and one [s] when a [coronal] consonant followed.

From the data it also becomes evident that there are clear asymmetries in the patterns of assimilation. [Coronal] sounds assimilate more frequently (7.8%) than other places of articulation; cf. [dorsal] (1.5%) and [labial] (3.7%).[11] Another asymmetry concerns the manner of articulation of the targets that undergo assimilation. Nasal sounds are more prone to assimilation than stops, and fricatives assimilate the least.

The question we turn to now is whether these results are special to function words or whether they form a general pattern observable in connected speech.

2. Lexical words

For lexical words, we counted a total of 2916 possible environments for regressive place assimilation. As compared to function words, there were more C_1 [dorsal] segments. Of all possible environments, 127 (4.4%) assimilations were actually realized. An overview over the different targets and triggers is presented in Table III(a)–(c).

The data for lexical words follow a similar assimilation pattern to that of the function words. [Coronal] segments undergo regressive place assimilation in 121 cases, of which 97 were nasals [Table III(b)]. Among the nasals, 8 /n/ (7.4%) were realized as [ŋ]. The rest, i.e., 89 /n/ (9.4%) were produced as [m]. For lexical words, final [t]s accounted for 24 cases (4.5%) of regressive assimilations. Of the 24 instances where [t] was assimilated, there was one utterance where [t] became [k] (1.8%), 23 cases showed assimilation to [p] (4.9%). No [coronal] fricative changed place of articulation. As for [labial] target segments, there occurred one assimilation: a word final [m] assimilated to [n] preceding a coronal stop [Table III(a)]. No other labial segment assimilated. [Dorsal] segments assimilated 5 times, all of them were [x]; 3 of them assimilated to [labial], 2 to [coronal] [Table III(c)].

Overall, the data of the lexical words also revealed two kinds of asymmetries. First, the nasal consonants assimilated more often than stops or fricatives. The second asymmetry concerns again the place of articulation of the target segment: [coronal] sounds undergo regressive place assimilation much more frequently (6.1%) than [labial] (1.3%) or [dorsal] (0.6%) segments.

C. Comparison of function and lexical words

The general pattern of assimilation was the same for lexical and function words although the latter underwent assimilation more frequently. Overall, we analyzed 7060 possible sequences for regressive place assimilation in the Kiel corpus of which 393 instances of assimilation occurred (see Table IV); i.e., 5.6% of the possible sequences were actually assimilated. Function words assimilated in 266 cases, whereas lexical words account for 127 instances. However, we should also note that function words occurred more often as targets in our sample than lexical words. We counted 4144 function words (58.7%) and 2916 lexical words (41.3%) as targets; hence, 6.4% of the function words and 4.4% of the lexical words assimilated. Nonetheless, function words show a significantly higher degree of regressive assimilation, as a *Chi-Square* test revealed ($\chi^2=13.9$, $p<0.001$).

Figure 1 depicts the relative percentage of assimilations depending on manner and place of articulation of C_1 in lexical and function words. For both function and lexical words, nasals are the most frequent to assimilate (350 out of 393—89.1%). Overall, stops assimilate in 28 (7.1%) cases and fri-

TABLE II. C_1 Targets and C_2 Triggers for all assimilated function words. The lightly shaded cells highlight assimilations. (a) Function words ending in a [labial]. (b) Function words ending in a [coronal]. (c) Function words ending in a [dorsal].

(a) Function words ending in a [LABIAL]

Assimilation	LABIAL>CORONAL						LABIAL>DORSAL	
C_2 Triggers / C_1 Target / PLACE [LABIAL]		n	t,d	ts	z	ʃ		k,g
/m/ 27/583 4.6%	m>n 4.3%	1/76	13/204	4/79	5/177	0/4	m>ŋ 9.3%	4/43
/p, b, f, v/ 0/141	p>t	0/18	0/81	0/10	0/32	0/0	p>k	0/0
Sum 27/724 3.7%	23 / 681 3.4%						4 / 43 9.3%	

(b) Function words ending in a [CORONAL]

Assimilation	CORONAL > LABIAL					CORONAL > DORSAL	
C_2 Triggers / C_1 Target / PLACE [CORONAL]		m	p,b	pf	f,v		k,g
/n/ 225 / 1230 18.3%	n>m 16.2%	44 / 187	33 / 142	3 / 4	88 / 703	n>ŋ 29.4%	57 / 194
/t, d/ 4 / 200 2.0%	t>p 2.3%	4 / 43	0 / 21	0 / 2	0 / 107	t>k	0 / 27
/s/ 1 / 1021 0.1%	s>f 0.1%	0 / 138	0 / 161	0 / 1	1 / 534	s>x	0 / 187
/ç/ 2 / 510 0.4%	ç>f 0.6%	0 / 91	0 / 56	0 / 0	2 / 186	ç>x	0 / 177
Sum 232 / 2961 7.8%	175 / 2376 7.4%					57 / 585 9.7%	

(c) Function words ending in a [DORSAL]

Assimilation	DORSAL > LABIAL					DORSAL > CORONAL					
C_2 Triggers / C_1 Target / PLACE [DORSAL]		m	p,b	pf	f,v		n	t,d	ts	z	ʃ
/x/ 7 / 459 1.5%	x>f 2.8%	0/70	0/49	0/5	6/94	x>s 0.4%	0/51	0/98	0/24	1/37	0/31
Sum 7 / 459 1.5%	6 / 218 2.8%					1 / 241 0.4%					

TABLE III. C_1 target and C_2 triggers of all assimilated lexical words. The lightly shaded cells highlight assimilations. (a) Lexical words ending in [labial]. (b) Lexical words ending in [coronal]. (c) Lexical words ending in [dorsal].

(a) Lexical words ending in [LABIAL]

Assimilation	LABIAL>CORONAL						LABIAL>DORSAL	
C₂ Trigger / C_1 Target PLACE [LABIAL]		n	t,d	ts	z	ʃ		k,g
/m/ 1 / 34 2.9%	m>n 3.3%	0/5	1/20	0/2	0/2	0/1		0/4
/p, b, f, v/ 0 / 42		0/5	0/26	0/0	0/3	0/2		0/6
Sum 1 / 76 1.3%	1 / 66 1.5%						0 / 10 0%	

(b): Lexical words ending in [CORONAL]

Assimilation	CORONAL>LABIAL					CORONAL>DORSAL	
C_2 Trigger / C_1 Target PLACE [CORONAL]		m	p,b	pf	f,v		k,g
/n/ 97 / 1050 9.2%	n>m 9.4%	14/160	29/285	0/0	46/497	n>ŋ 7.4%	8/108
/t, d/ 24 / 531 4.5%	t>p 4.9%	18/146	3/143	0/1	2/184	t>k 1.8%	1/57
/s, ç, ʃ/ 0 / 386		0/68	0/59	0/0	0/165		0/97
Sum 121 / 1970 6.1%	112 / 1708 6.6%					9 / 262 3.4%	

(c): Lexical words ending in [DORSAL]

Assimilation	DORSAL>LABIAL				DORSAL>CORONAL					
C_2 Triggers / C_1 Target PLACE [DORSAL]		m	p,b	f,v		n	t,d	ts,tʃ	z	ʃ
/ŋ, k, g/ 0 / 342		0/25	0/12	0/78		0/27	0/158	0/13	0/22	0/7
/x/ 5 / 528 0.9%	x>f 2.5%	0/32	0/37	3/53	x>s 0.5%	0/15	0/343	0/18	2/26	0/4
Sum 5 / 870 0.6%	2 / 237 1.3%				2 / 633 0.3%					

catives in 15 instances (3.8%). Of a total of 393 assimilated targets, overwhelmingly the [coronal] sounds (353 out of 393—89.8%) assimilate to the place of a following segment across word boundaries, whereas [labial] (7.1%) and [dorsal] (3.1%) segments usually do not. In general, coronal targets (C_1) by far outnumber the other places of articulation (4931 or 69.8%). The fewest number of targets are [labial] sounds

TABLE IV. Assimilation of function and lexical words combined.

C_1 target				C_2 trigger					
Place	Total	Assimilated		[Labial]		[Coronal]		[Dorsal]	
[Labial]	800	28	3.5%	—		24/747	3.2%	4/53	7.5%
[Coronal]	4931	353	7.2%	287/4084	7.0%	—		66/847	7.8%
[Dorsal]	1329	12	0.9%	9/455	2.0%	3/874	0.3%	—	
Sum	7060	393	5.6%	296/4539	6.6%	27/1621	1.7%	70/900	7.8%

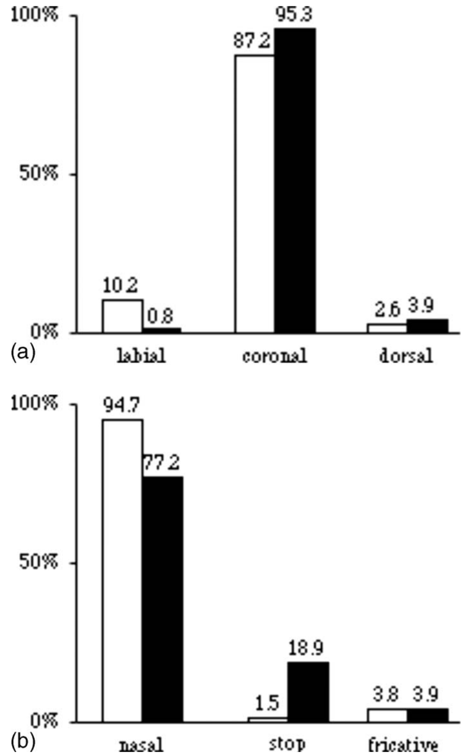

FIG. 1. Relative percentages of regressive place assimilations (based on the total number of assimilated sequences) across different place (a) and manner of articulation (b). Function words are represented by light bars, lexical words by dark bars.

(800 or 11.3%). The only [dorsal] segment—both in function words as well as lexical words—that assimilates is [x].

Note that the analysis did not differentiate between C_1 and C_2 sequences that were within one phrase or sequences that crossed phrase boundaries. Out of the 7060 items analyzed in the data, there were 1174 (16.6%) crossing either a period, a question mark, or a comma in the transcription. Of all 18 cases where C_1 and C_2 were separated by a question boundary, none showed assimilation. Concerning periods, there were overall 310 sequences in this category. There was one (out of 188) assimilation occurring in a [coronal]-[labial] context. Overall a comma separated 848 of the 1174 sequences. In this category, there were 13 assimilations. 10 (out of 319—3.1%) occurred in a [coronal]-[labial] context, 2 (out of 42—2.4%) showed an assimilation of [x] to [f] in front of [f], and two cases (out of 441 possible sequences—0.5%) had an assimilation of [x] to [s] in front of [z]. Thus, although phrase boundaries do impede assimilation, at least for commas, there are cases where assimilation even occurs across those boundaries.

To summarize, across word place assimilations in German is controversial. Some authors claim that such assimilations do not occur (cf. Wurzel, 1970; Vater, 1979; Wiese, 1996), while others assert the opposite (cf. Kohler, 1995). This controversy led us to systematically analyze assimilations across word boundaries in conversational German. The Kiel corpus data suggests that although such assimilations are not frequent, they do occur—overall, approximately 6% of possible assimilatory sequences did undergo a change in place of articulation. We analyzed function and lexical words separately since they are claimed to be different, and indeed, we found a significant difference in the number of assimilations between the two categories although the assimilation patterns were the same. Function words are more likely to assimilate than lexical words. Moreover, the data revealed clear asymmetries in the pattern of assimilations that actually occurred. One asymmetry concerns the place of articulation of the targets (C_1) undergoing assimilation: [coronal] sounds are more frequently assimilated than [dorsal] and [labial] consonants. A second asymmetry is that nasals assimilate more often than stops or fricatives.

The assimilation data we have analyzed and presented are based on the transcriptions of trained phoneticians who noted sequences where assimilation had occurred despite the fact that they had the orthographic as well as the canonical phonetic transcription that could have biased them to perceive the canonical sounds. Speech perception research suggests, however, that although some sounds might seem to be assimilated, there may still be residual cues for listeners to identify the underlying segments (cf. Gow, 2002). We therefore turn to the perception of naive listeners and compare them to the transcriptions of the trained phoneticians.

IV. IDENTIFICATION OF TRANSCRIBED PLACE OF ARTICULATION

The question we investigate next is: Do naive listeners (naive both with respect to the goal of the experiment as well as not having additional information from the context) perceive the assimilated and unassimilated segments from the Kiel dialogues in the same way as trained phoneticians who used speech analyses tools? We conducted two phoneme identification experiments—a forced choice and a free choice task—using material from the Kiel corpus. We opted to focus on nasals (/n/ and /m/) since the choice of assimilated segments was larger than for oral stops and we were able to take stimuli from several speakers thereby lessening speaker dependence (for details see Sec. IV A 1 below).

A. Experiment 1: Phoneme identification

A timed forced-choice identification task was chosen for the first experiment. Subjects had to decide whether the auditory stimuli included either a labial [m] or a coronal [n]. This method was chosen to determine the speed as well as the accuracy of the subjects' decision. Other studies have shown that assimilations can be only partial and that listeners are sensitive to residual cues left (cf. Nolan, 1992; Manuel, 1995; Gow, 2002). Manuel (1995), for example, found that in a sequence [nð] in *win those*, where the /ð/ became a nasal, the place of articulation was not that of a "real" [n], suggesting that some featural information was still available to the listener. Our focus was not just on the assimilated stimuli, but also stimuli that had been labeled as unchanged from the canonical—that is underlying /n/ or /m/ which were spoken and heard as [n] and [m]. The issue was whether the responses to the unchanged stimuli differed across varying contexts—vowel, labial, dorsal, coronal. The crucial conditions with a set of examples are listed in Table V. The segmental context from which the stimuli were extracted is

TABLE V. Examples of stimuli with the vowel [e:] for experiments 1 and 2. Column 1 gives the Kiel transcription. Column 2 provides the orthographic contexts from which the stimuli were extracted and column 3 gives the three conditions—unchanged unassimilated-/m/ and unassimilated-/n/, and assimilated.

Kiel corpus transcription	Example stimuli in orthography	Condition
		Unassimilated-/m/
[e:m]	…von d**em** achtzehnten Juni?…	Vowel context /m/-vowel
[e:m]	…mit d**em** Bericht…	Labial context /m/-labial
[e:m]	…dann d**em** Dienstag…	Coronal context /m/-coronal
[e:m]	…und d**em** ganzen Kram…	Dorsal context /m/-dorsal
		Unassimilated-/n/
[e:n]	…Freitag, d**en** ersten…	Vowel context /n/-vowel
[e:n]	…für d**en** Bericht…	Labial context /n/-labial
[e:n]	…in d**en** deutschen…	Coronal context /n/-coronal
[e:n]	…d**en** ganzen Tag…	Dorsal context /n/-dorsal
		Assimilated
[e:m]	…über d**en** Bericht…	Labial context

double underlined. Since the coronal nasal assimilated most frequently, we only used conditions where /n/ was assimilated to [m].

Our predictions are the following. Most descriptions of assimilations suggest that coronal consonants are more vulnerable to variation in the context of consonants with other places of articulation (cf. Paradis and Prunet, 1991). Consequently, one could expect that labial and dorsal C_2 contexts would leave more acoustic traces in unassimilated-/n/ stimuli than coronal and dorsal C_2 segments influence unassimilated-/m/ stimuli. This would make it more difficult for listeners to come to a definite decision for the unassimilated-/n/ stimuli. Therefore, we expect slower reaction times (RTs) for unassimilated-/n/ in labial and dorsal contexts but no reaction time differences for those items in the vowel or homorganic consonantal context. Insofar as the difference between assimilated-/n/s and unassimilated-/m/s are concerned, we expect no difference in the speed of reaction, assuming that the assimilated-/n/s exhibit complete neutralization. However, whether the assimilated-/n/ items were equally well heard as [m] as the unassimilated-/m/s depends on whether the assimilation as perceived by the transcribers was reasonably complete. Thus, both the reaction time measures as well as percentage of [m] and [n] responses are vital.

1. Materials

The stimuli for the perception task consisted of a vowel-nasal (VN) sequence extracted from real words (CVN or VN), and were taken from 27 different speakers (13 female, 14 male) of the Kiel corpus. At most five items were taken from any given speaker. We thereby kept the segmental context as similar as possible and at the same time were able to make the perception task speaker-independent. The two vowels in the VN sequences we chose were transcribed as either a mid [eː] or a low [ɑ] vowel. The extracted sequences with [ɑ] form possible words: *an* [ɑn] "on, at.acc" and *am* [ɑm] "at.dat," whereas the [eːn] and [eːm] sequences do not. A set of sentences from which the [eː] sequences were extracted is given in Table V and corresponding [ɑm/ɑn] sequences are given in Appendix B.

We cut the VN-items at zero-crossings in order to avoid clicks at item boundaries using both visual as well as auditory information. The first identifiable glottal period was taken as the beginning of the vowel. However, when there was an extensive amount of coarticulation from the preceding segment (i.e., at the word onset), we cut off up to four glottal periods to ensure that the consonantal onset could no longer be perceived. The end of the nasal in the VN-items was determined when the amplitude of the waveform dropped markedly or at the beginning of the closure of the following consonant. Thus, the nasal itself was left untouched, but any contextual information in the following closure would have been removed.

For each vowel (i.e., [eː]/[ɑ]), we chose ten [coronal]#[labial] assimilated sequences (assimilated category), and 10 each of unassimilated [coronal] (unassimilated-/n/) and [labial] (unassimilated-/m/) items. This added up to 60 different stimuli. The unassimilated items were cut out of different contexts (see Table V and Appendix B); three preceded a [labial] consonant, three a [coronal] consonant, two a [dorsal] consonant, and two were originally followed by a vowel. The amplitude of the items was equalized.

2. Subjects and procedure

Overall, 18 undergraduates from the University of Konstanz with no reported hearing disorders served as subjects in the experiment and were paid for their participation. They were tested in groups of 5 or less and were given oral as well as written instructions. A push-button box with two buttons labeled [m] and [n] was placed in front of each subject. They were instructed to listen to the syllables presented over headphones and decide as quickly as possible whether the consonant was [m] or [n] and press the appropriate button with the index finger of their dominant hand. Before the test began,

TABLE VI. Least square means of reaction times for three main categories in all contexts for both [m] and [n] responses [% values are computed for each row by $N_x/(N_{\text{Response[m]}}+N_{\text{Response[n]}})$].

	Response [m]			Response [n]		
Context	N	%	RT ms	N	%	RT ms
Unassimilated-/m/	1643	93.2	536.3	120	6.8	580.4
/m/-labial	467	89.1	535.1	57	10.9	518.3
/m/-coronal	523	97.9	531.0	11	2.1	573.2
/m/-dorsal	310	88.6	547.9	40	11.4	647.8
/m/-vowel	343	96.6	531.3	12	3.4	582.1
Unassimilated-/n/	405	23.1	547.1	1346	76.9	547.2
/n/-labial	141	26.8	592.9	385	73.2	553.4
/n/-coronal	92	17.6	520.6	432	82.4	528.4
/n/-dorsal	95	27.2	536.2	254	72.8	570.1
/n/-vowel	77	21.9	538.6	275	78.1	536.8
Assimilated (labial context)	1534	87.5	545.8	219	12.5	580.0

the subjects familiarized themselves to the task with practice items, but were given no feedback about the "congruency" of their decisions.[12]

Each item occurred five times during the experiment, adding up to 300 items presented in a randomized order. The sequence of presentation was as follows. Each item was preceded by a warning tone of 300 ms followed by 200 ms of silence. After each test stimulus, there was a pause of 1500 ms where subjects had time to push the button and the next sequence began. Reaction time measurements began at the onset of the nasal segment. The stimuli were played from a SONY DAT recorder and presented over headphones (Sennheiser HD520II). In the setup, a central experimental hardware box connected the DAT recorder, the response boxes and a Macintosh computer, where the reaction times were recorded (Reetz and Kleinmann, 2003). A single experimental session lasted approximately 18 min excluding the practice items.

3. Results

The responses of all 18 subjects went into the reaction time analysis.[13] Responses faster than 200 ms and slower than 1000 ms were disregarded leading to the exclusion of 133 responses (2.5% of the data). None of the subjects showed an exceptionally high number of responses which were too slow or too fast. Reaction times as a dependent variable and the factors *subject* (as random), *response* ([m] or [n]),[14] *underlying* (unassimilated-/m/, unassimilated-/n/, assimilated), *context* (nested under *underlying*) (/n/-coronal, /n/-labial, /n/-dorsal, /n/-V, /m/-coronal, /m/-labial, /m/-dorsal, /m/-V, assimilated), *item* (nested under *underlying* and *context*), *response* × *context* (nested under underlying) and *underlying* × *response* as independent variables were entered into an analysis of variance (ANOVA) with REML estimation.[15] There was a main effect of *context* [$F(65\,181)=9.03$, $p<0.001$] and *response* [$F(15\,181)=15.37$, $p<0.001$], and the interaction of *context* × *response* was also significant [$F(65\,181)=4.70$, $p<0.001$]. *Speaker* and *repetition* were not significant factors in the ANOVA. They are therefore not reported. The least square means of the RT measures for both [m] and [n] responses for each *context* are given in Table VI.

Several pairwise posthoc comparisons were made for the critical conditions, the interpretations of which are summarized below with individual figures.

(i) Recall that based on the analysis of the Kiel corpora transcriptions, the expected congruent responses are [m], for the unassimilated-/m/ category and [n] for the unassimilated-/n/ category. The percentage congruent responses are revealing. For unassimilated-/m/ stimuli, 93% of the responses were [m], and only 7% were [n]. In contrast, for unassimilated-/n/ items, almost a quarter of the stimuli were identified as the opposite [m]—77% [n] versus 23% [m]. Obviously, listeners had more difficulty with the unassimilated-/n/ stimuli than with unassimilated-/m/ items. A *Chi-Square* analysis revealed a significant difference ($\chi^2=1773.63$, $p<0.001$). The reaction times also reflect the same pattern. If we consider the congruent responses, [m] for unassimilated-/m/ and [n] for the unassimilated-/n/, the reaction times across these categories (536 and 547 ms, respectively) are statistically different ($t=2.15, p<0.05$). There is a much larger difference between the reaction times for [m]- and [n]-responses to the unassimilated-/m/ stimuli (536 ms versus 580 ms, $t=2.97$, $p<0.05$). Likewise, there is a significant difference between the incongruent [m]-responses of unassimilated-/n/ and the [n]-responses of unassimilated-/m/ (547 ms versus 580 ms, $t=2.04$, $p<0.05$). The RT of [m] or [n] responses to the unassimilated-/n/ category are essentially identical. This suggests that it was more difficult for the listeners, and hence, they were slower, to give [n] responses to unassimilated-/m/ stimuli when they were uncertain.

(ii) Since there were four contexts, the next point to address is if any particular context is responsible for the worse identification of unassimilated-/n/ than

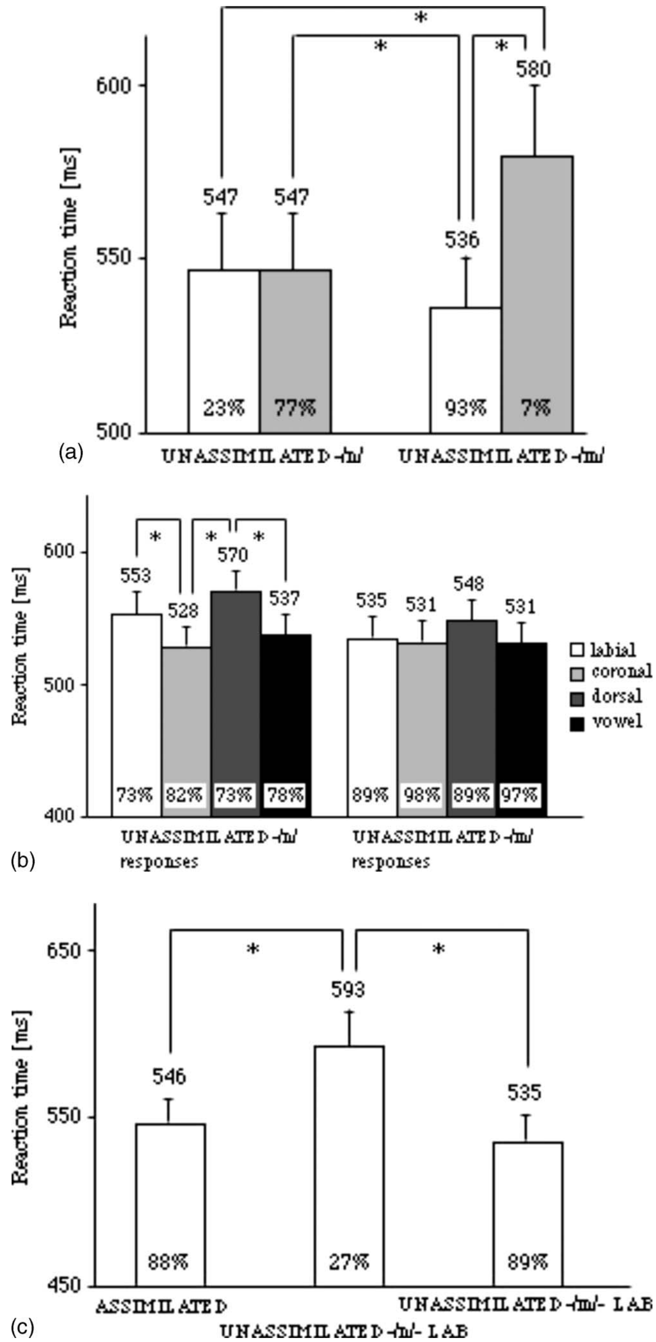

FIG. 2. (a) [m] and [n] responses to unassimilated-/n/ and unassimilated-/m/ stimuli in percent and with their reaction times as bars. Asterisks indicate significant differences in reaction times. White bars represent [m]-responses and gray bars show [n]-responses. (b) [n] responses to unassimilated-/n/ and [m] responses to unassimilated-/m/ stimuli differentiated by context. Percent of responses are given in numbers and the bars represent the reaction times with significant differences indicated by asterisks. (c) Percentages and reaction times for [m] responses to assimilated, unassimilated-/n/-labial, and unassimilated-/m/-labial stimuli. Asterisks indicate significant reaction time differences.

unassimilated-/m/ [see Fig. 2(b)]. With respect to percentage congruent responses, in all contexts more than 89% of the unassimilated-/m/ stimuli were congruently responded to as [m]. This was not so for the unassimilated-/n/ stimuli, where 27% of the responses were [m] in the labial and dorsal contexts. When an unassimilated-/n/ item was preceding another coronal or a vowel, the responses were more comparable to the unassimilated-/m/ stimuli, viz. around 80% [n] responses. To test whether parallel results are reflected in the reaction times, we ran pairwise comparisons across all four contexts—vowel, coronal, dorsal, labial [see Fig. 2(b)]. For the [m] responses to unassimilated-/m/, there were no significant differences in reaction across any of the contexts.

Thus, unassimilated-/m/ (extracted from *dem, am*, etc.) stimuli were heard and reacted to as [m] equally fast regardless of which context they had been extracted from. Would we find the same pattern for [n] responses to the unassimilated-/n/ category stimuli? Based in the corpus analysis, we know that /n/ is more vulnerable to coarticulation from following consonants with different places of articulation. There could therefore be a difference between the contexts dorsal, labial, on the one hand, versus coronal and vowel, on the other. In the former contexts, the /n/ may have more coarticulation cues of the place of articulation of the following dorsal or labial consonant, making it more difficult to label the unassimilated-/n/ as [n] in a reaction time task, whereas in the coronal context, the /n/ is in its ideal environment. The pairwise comparisons confirmed this prediction. The [n] responses to unassimilated-/n/ in coronal context differed significantly from the responses to unassimilated-/n/ in labial context ($t=-2.82$, $p<0.005$) as well as from the dorsal contexts ($t=-3.99$, $p<0.001$). Another significant difference emerged in the comparison of the [n]-responses to unassimilated-/n/ in the dorsal and the vowel contexts ($t=-2.91$, $p<0.005$). There were no further significant differences between any other contexts for the [n]-responses. Thus, the [n]-responses to unassimilated-/n/ in the coronal and vowel contexts, which are the most neutral contexts in terms of coarticulation, are significantly different from the labial and dorsal contexts, We can therefore conclude that the coarticulation cues from the (deleted) following labial and dorsal consonants were strong enough to slow down the subjects' [n] responses to these stimuli. Recall that these consonants had been labeled as [n] by phoneticians who had recourse to both visual and auditory cues and were under no time pressure.

In sum, the labial and dorsal contexts had a slowing down effect on the [n] responses for unassimilated-/n/ stimuli as compared to its homorganic coronal context. This effect is not observed for the unassimilated-/m/ stimuli in the coronal and dorsal contexts in comparison to its homorganic labial context. For the unassimilated-/m/ stimuli, the subjects' speed and their response were unaffected by the context of other places of articulation, from which we may deduce that there were less coarticulation cues which could confuse them. Thus, there was an asymmetry in the stimuli even where trained phoneticians had transcribed the sounds carefully.

(iii) The assimilated stimuli were always (by definition) extracted from a labial context. The crucial question to gain further insight in how far the assimilations were produced completely is whether these stimuli differ from the unassimilated-/m/ stimuli in the same context. The unassimilated-/m/ stimuli in labial context can be seen as the most prototypically produced labial features without coarticulation and they are taken as clear examples of [m]. Since we found an effect of coarticulation of the labial context in the unassimilated-/n/ stimuli, we also take these for comparison. With respect to percent congruent responses, the [m] responses to the assimilated stimuli and the unassimilated-/m/-labial were almost identical—88% versus 89%. Further, there were no significant differences in reaction times in the [m] or [n] responses to these categories. From this we can conclude that subjects were equally fast in responding to the assimilated [m] and the canonical /m/ stimuli (e.g., [e:m] from *über de**n** **B**ericht* versus [e:m] from *mit de**m** **B**ericht*).

As for the reaction time of [m] responses to unassimilated-/n/-labial stimuli, these were different from the [m] responses to the other two categories (assimilated versus unassimilated-/n/-labial $t=-4.08$, $p<0.001$; unassimilated-/n/-labial versus unassimilated-/m/-labial $t=-4.64$, $p<0.001$), indicating that although there was sufficient coarticulation, these stimuli were different from those that were considered by the transcribers as real assimilated or canonical unassimilated-/m/ items. Crucially, there is no difference between the [m] responses in the assimilated and the unassimilated-/m/-labial categories ($t=-1.65$, $p<0.1$). Thus, for listeners, the assimilated stimuli were similar to the unassimilated-/m/-labial but not to the unassimilated-/n/-labial.

Recall that the task in experiment 1 was forced choice where subjects had to choose between [m] or [n] as possible responses. To determine in how far the forced choice task of experiment 1 created a possible bias in the subjects' responses, we ran a second experiment where the listeners were free to choose and write down what they heard.[16] Our hypothesis was that if we obtained the same pattern of results, then we could conclude that the context-dependent responses of unassimilated-/n/ stimuli was caused by the fact that the listeners were forced to choose between [m] or [n]. Further, we also wanted to examine the pattern of responses to the dorsal stimuli since in experiment 1 the listeners had no option of providing dorsal responses.

B. Experiment 2: Phoneme transcription task

1. Material and design

In order to analyze if possible confounds in the forced-choice task could have affected the results, we used a phoneme-identification task where subjects could write down in a booklet what fragment they heard. The stimuli were identical to experiment 1 except that there was a longer pause between two items (2500 ms instead of 1500 ms), sufficient for writing the syllables but not too much time to think about the stimuli. Each page in the booklet had space for ten items. Warning tones were added after every ten

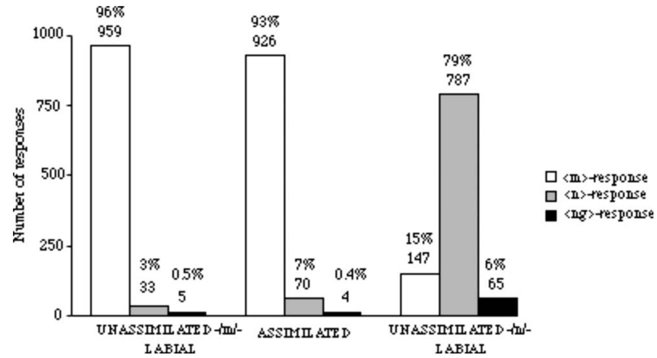

FIG. 3. Total number of responses and percentages within the three main categories.

items, prompting subjects to turn to the next page of the booklet. This was done to ensure that if a subject missed an item, it was possible to correctly resume at the beginning of the next page. Thus, as in experiment 1, subjects listened to 300 stimuli.

2. Subjects and procedure

Ten students from the University of Konstanz served as subjects, and none had taken part in the earlier experiment. They were tested individually and were paid for their participation. The set up and equipment was the same as in experiment 1. Written instructions were given to the subjects prior to the experiment and they received the same practice items as before. They were asked to write down what they heard as quickly and accurately as possible. No instruction was given with reference to nasals, syllables or the "wordness" of the items. Given German orthography, if subjects heard nasals, we expected subjects to transcribe them using one of the three possible responses ⟨m⟩, ⟨n⟩, or ⟨ng⟩.

3. Results

In all, there was only one missing response and three were not a nasal. These four items were discarded (0.13%). The nasal responses were split up into the three main categories as above (labial, assimilated, and coronal), based on the original labeling in the Kiel corpus. A total of 2996 transcribed items went into the analysis. Across all categories subjects heard 2032 ⟨m⟩ (67.8%), 890 ⟨n⟩ (29.7%), and only 74 ⟨ng⟩ (2.5%), of which 41 (i.e., 55.4%) come from unassimilated-/n/ in a dorsal context.

Within the individual categories, the nasal segments were transcribed as follows (see Fig. 3). Unassimilated-/m/ segments were transcribed as ⟨m⟩ in 959 cases (96.2%), ⟨n⟩ in 33 instances (3.3%), and ⟨ng⟩ in 5 (0.5%) cases. Assimilated tokens were transcribed as ⟨m⟩ in 926 cases (92.6%), as ⟨n⟩ in 70 (7.0%) cases, as ⟨ng⟩ in four cases (0.4%). Unassimilated-/n/ were transcribed as ⟨n⟩ in 787 (78.8%) cases, ⟨m⟩ 147 times (14.7%), and ⟨ng⟩ in 65 (6.5%) instances.

Insofar as congruent responses are concerned, the percentage of [m] responses to assimilated and unassimilated-/m/ categories is far higher than the corresponding [n] responses to the unassimilated-/n/ category (93%, 96%, versus 79%).

4. Discussion

The free choice task was taken on to ensure that the incongruent responses in experiment 1 were not due to the fact that subjects were forced to choose between two nasals. In particular, we were concerned that the large number of ⟨m⟩ responses to unassimilated-/n/ stimuli was biased by the forced choice task. However, experiment 2 shows that this was not the case. First, there were only three non-nasal responses, and second, 97.6% of the entire responses were transcribed as ⟨m⟩ or ⟨n⟩.

In fact the pattern of results was the same as in experiment 1. On the whole, the unassimilated-/n/ stimuli were more difficult to identify congruently as ⟨n⟩ (79%) and were subject to context dependent responses, as compared to the unassimilated-/m/ or assimilated items, both of which were congruently identified as ⟨m⟩, 96% and 93%, respectively. As in experiment 1, the unassimilated-/n/ stimuli in the context of labial consonants were identified as ⟨m⟩ 15% of the time (experiment 1: 27%). In contrast, there were only 3% ⟨n⟩ responses to unassimilated-/m/ items. Overall, the accuracy of experiment 1 for [labial] and assimilated tokens was even higher in experiment 2, possibly due to the longer time subjects had for their decisions. The results for the assimilated category are very much the same as in experiment 1. They were largely perceived as [labial], indicating the completeness of assimilation. In general, this experiment replicates the same asymmetry we observed already in the identification task and the corpora analysis. One remaining issue is the acoustic differences between the different conditions, critically between the assimilated labial, the canonical unassimilated-/m/ against the unassimilated-/n/-coronal. Since the assimilated nasals did not differ in perception from the canonical unassimilated-/m/, one would conjecture that the acoustic differences would also be minimal.

C. Acoustic measurements

One important issue that has also been reflected in the literature on place assimilation is the question whether acoustic cues can be found that relate to listeners' decisions for [n] or [m] (e.g., Nolan, 1992; Gow, 2002; Dilley and Pitt, 2007). Following Dilley and Pitt's (2007) approach, we investigated the stimuli from the experiments. In their study, they compared assimilated segments with their underlying counterparts. Since their results are based on the variation in the F2 of the preceding vowel, we took the same measure and applied it to items from our perception test. We opted for the analysis of the most prototypical items. Therefore, we compared the assimilated items with unassimilated-/m/ stimuli in labial context and unassimilated-/n/ stimuli in coronal context. Since the number of items from the experiments was too small for calculating an ANOVA, we randomly selected additional items from the Kiel corpus.

There is one important difference between our stimuli compared to Dilley and Pitt's (2007): In our stimuli, the final consonant, in our case the nasal, was not deleted and acoustic information on place of articulation can also be extracted from the nasal segment, therefore, we also took F2 measurements at the midpoint of the nasal segment.

1. Method

We measured the difference in the F2 frequency values in hertz between the middle of the vowel and immediately before the beginning of the nasal murmur of all 20 assimilated items, six unassimilated-/m/-labial, and six unassimilated-/n/-coronal items that were used in the perception studies as an indication for the amount of possible assimilation. In order to base a statistical analysis on a more thorough database, we randomly added the measurements of four assimilated and 18 unassimilated-/m/-labial and 18 unassimilated-/n/-coronal items with the respective vowels. Overall, the measurements of 72 items were analyzed—36 for each vowel (i.e., [eː]/[ɑ]), 24 for each condition (i.e., assimilated, unassimilated-/m/-labial, and unassimilated-/n/-coronal). As in Dilley and Pitt, a mixture of automatic and hand taken measurements was performed (Dilley and Pitt, 2007). Formant values were taken form the estimation provided by PRAAT (Vers. 4.6.04, Boersma & Weening) and wide-band spectrograms. In case that the estimated formant values differed from the spectrograms we followed the spectrograms readings. Dilley and Pitt could measure only the difference between midpoint and endpoint of vowels to gain information about the place of articulation of the upcoming segments, since their items included cases where the consonant in question had been deleted. Since the nasal consonant was never deleted in our case, we were able to measure the F2 frequency in the midpoint of the nasal segments (F2 measurements on the nasal differ for labial and coronal nasal consonants, cf. Stevens, 1998, pp. 487–507). The F2 values at the midpoint of the nasals were measured the same way as in the vowels.

2. Results

F2 differences in the midpoint and endpoint of preceding vowels were subject to an ANOVA with *condition* (assimilated, unassimilated-/m/-labial, and unassimilated-/n/-coronal) and *vowel* as independent variable, as well as the interaction of the two factors (*vowel*$\times$*condition*). Post-hoc tests were performed for the contrasts between the *conditions*. Figure 4(a) summarizes the results for the F2 differences.

As for the F2 difference analysis, there was a main effect of both *condition* $[F(2,66)=10.7106,\ p<0.002]$ and *vowel* $[F(1,66)=3.3052,\ p<0.05]$, but no significant interaction. A post hoc test revealed that unassimilated-/n/-coronal items were significantly different from assimilated ($t=-2.317$, $p<0.05$) and unassimilated-/m/-labial ($t=2.1242$, $p<0.05$) items, but the latter two were not significantly different from each other.

For the F2 measurements taken at the midpoint of the nasal consonants [see Fig. 4(b)] we used the same ANOVA design and gained the following results: There was a main effect of *condition* $[F(2,66)=5.1775,\ p<0.01]$, but no effect of *vowel*, and no interaction. A post-hoc test showed that

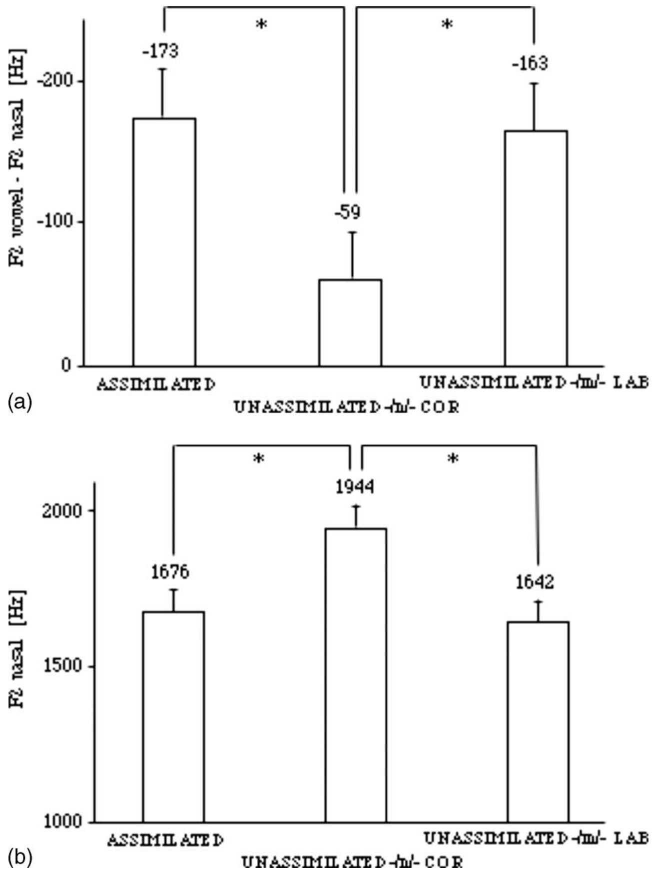

FIG. 4. (a) Differences between F2-frequency measures at the middle and at the end of the vowel in hertz. Significant differences between these difference values are marked by asterisks. (b) Least square means of F2-frequencies at nasal midpoints for the investigated conditions, significant differences are indicated with asterisks.

unassimilated-/n/-coronal items were significantly different from assimilated ($t=-2.605$, $p<0.02$) and unassimilated-/m/-labial ($t=2.9385$, $p<0.005$) items, but the latter two were not significantly different from each other. Figure 4(b) depicts the least square means of the nasal F2 measurements.

3. Discussion

The F2 differences between vowel and nasals and the F2 measurements on the nasals correspond to our perception results. There is no significant difference between the F2 of the assimilated coronals and the canonical unassimilated-/m/. Nor does the nasal F2 differ in these two categories. Corresponding to the perception results, there is a significant difference both in the F2 of the nasal and the F2 difference for the assimilated and unassimilated-/m/-labial nasals on the one hand and the unassimilated-/n/-coronal on the other. The results indicate that subjects take these acoustic cues as basis for their decision when deciding on whether they heard [m] or [n].

V. GENERAL DISCUSSION

The focus of this paper was to investigate the extent to which regressive place assimilations across words exist in conversational German and how listeners perceive them. Analyzing the Kiel corpus of spontaneous speech (IPDS, 1994; 42 speakers, approximately 4 h of recordings) we found that place assimilation does occur across words in German in approximately 6% of all possible sequences of consonants differing in place of articulation (cf. Sec. III C). This is slightly less than the percentage of assimilation reported by Dilley and Pitt (2007) for American English. We based our conclusions on the transcriptions made by trained phoneticians who had recourse to the speech signal as well as the context. Function words were more likely to assimilate than lexical words. Moreover, there was an asymmetry in the direction of assimilation. According to the transcriptions, coronal sounds [t, s, ʃ, ç, n] were more likely to assimilate than labial [p, f, m] or dorsals [k, x, ŋ] (see Table IV). A further asymmetry concerned the manner of articulation of consonants which were more likely to assimilate; nasal consonants were far more likely to assimilate than obstruents.

Exploiting the asymmetries, we ran two perception experiments (forced choice and free choice) to test how fast and accurately naive listeners' responses would correlate with the transcriptions. The two perception experiments using stimuli (from 27 speakers) labeled in the corpora as assimilated (/n/ > [m]) or unassimilated (unassimilated-/n/, unassimilated-/m/), showed that the Kiel transcription is very accurate concerning regressive place assimilations. There is a high correlation between the phoneticians' transcriptions and the listeners' judgments. Thus, when the subjects had to decide whether they heard a [coronal] or a [labial] nasal in a VN-syllable without additional information from context, they were very accurate for the assimilated and the unassimilated-/m/ stimuli. The assimilated sounds were overwhelmingly transcribed as [m] both in a speeded forced choice task ([m] or [n]) as well as in a free identification task. In the identical labial context, the reaction times for unassimilated-/m/ and assimilated also did not differ [see Fig. 2(c)]. That is, subjects were equally fast in making an [m] response to stimuli which were really /m/ and those that had been labeled as assimilated from /n/ to [m] in a labial context.

The only nonequivalence between the transcriptions and the listeners' responses was the unassimilated-/n/ category in the context of labial or dorsal segments. If at all, the transcribers were conservative in their judgments of which sounds assimilated as is indicated by the high amount of variation in the /n/-category. As compared to unassimilated-/m/ which was congruently identified as [m] 93% of the time in experiment 1, only 77% of the unassimilated-/n/ stimuli were identified as [n]. The same pattern shows in experiment 2: 79% [n] responses to unassimilated-/n/ stimuli as compared to 96% [m] responses to unassimilated-/m/ stimuli and 93% [m] responses to the assimilated stimuli. The reaction times in experiment 1 also correspond to the accuracy data. The differences in the RT of congruent and incongruent responses of unassimilated-/m/ and unassimilated-/n/ responses are revealing. First, the incongruent [n] responses to unassimilated-/m/ stimuli are significantly slower than the corresponding incongruent [m] responses to unassimilated-/n/ stimuli, indicating that the former were less transparent for the listeners. Second, there is a stronger context effect for the unassimilated-/n/ stimuli than for the unassimilated-/m/ stimuli. For the latter, there were no differences in reaction to the congruent [m] responses. However, the reaction times for the congruent [n] responses to the unassimilated-/n/ stimuli differed by context. For instance, in the neutral vowel and identical coronal contexts, the reaction times were much faster than in the labial and dorsal contexts. One could argue that coarticulation effects of the following deleted segments cause the incongruent responses, but this does not explain the asymmetry between unassimilated-/m/ and unassimilated-/n/ stimuli. That is, the coarticulation effects on the unassimilated-/n/ stimuli were presumably strong enough for listeners to respond to them with [m] as fast as they responded to the unassimilated-/n/ stimuli in their ideal coronal context. Almost a quarter of the unassimilated-/n/ stimuli were incongruently identified as [m] in both experiments. We see this as an indication that the transcribers of the Kiel corpus were "conservative" and labeled the unassimilated-/n/-labial as [n] rather than [m]. Since experiment 2 used a free-choice task, and the results were essentially the same to experiment 1, it is unlikely that subjects were forced to choose ⟨n⟩ as an answer as substitute for any other perceived segment.

We further investigated the acoustic patterns across the stimuli used in the perception experiments (Sec. IV C). Since our stimuli consisted of vowel-plus-nasal syllables, following Dilley and Pitt (2007) we took the F2 measures of the middle and end of the vowel, we also examined F2 at the nasal midpoint. Corresponding to the perception results, we found that the change in the F2 from the middle to the end of the vowel did not significantly differ between the unassimilated-/m/ and assimilated consonants. Similarly, the nasal formant measure did not differ between these categories indicating that the assimilated tokens shared these acoustic categories with the canonical /m/.

Both the perception results and acoustic analysis of the stimuli suggest that segments labeled as assimilated by the transcribers are indeed recognized not as tokens of the underlying words, but as perfect instances of the changed sound, i.e., complete assimilations do occur in running speech ($/n/ > [m]$ in a labial context). Actually, some assimilated tokens are judged by subjects in our experiments to be [m] 100% of the time. Clearly however, there exists gradience in the assimilation as we have seen in the response to the unassimilated-/n/-labial stimuli. Although transcribers labeled them as [n] they were often perceived as [m]. Gradualness of assimilation is most important for the [coronal]-category where we see the greatest amount of (response) variation.

The asymmetry between coronal versus dorsal and labials both in production analysis (coronal consonants assimilate more than the others) and in perception (coronals vary most in perception) has been frequently noted in the literature (cf. Lahiri and Evers, 1991; Paradis and Prunet, 1991; Ghini, 2001). The unmarkedness and asymmetry of coronals are tackled by markedness rules or other phonological principles (cf. Clements, 2001) or built into a recognition model with underspecification as in FUL (Lahiri and Reetz, 2002).

The results of the corpus study as well as the two experiments are fully compatible with the FUL model of speech perception (Lahiri and Reetz, 2002). The model assumes that segments can be specified with a [labial] or [dorsal] feature for place of articulation, but do not possess a [coronal] feature, that is, they are underspecified. Evidence for underspecification has been presented in different languages and for different phonological processes (among others, Lahiri and Evers, 1991; Lahiri and Marslen-Wilson, 1991; Ghini, 2001, Wheeldon and Waksler, 2004; Scharinger, 2006; cf. also Paradis and Prunet, 1991). Segments specified for features [labial] or [dorsal] lead to corresponding labial and dorsal places of articulation in production. Segments which are underspecified for place of articulation are produced by the default feature [coronal]. Moreover, since these are underlyingly not specified for place of articulation, they can assimilate more easily to the place of articulation of neighboring [labial] and [dorsal] segments, but not *vice versa*. There can be instances where segments are produced in-between two categories due to overlapping gestures. FUL predicts that if this assimilation takes place before the default production rule applies, assimilation can lead to a complete neutralization of the place of articulation contrast. This is what we find in the corpus study, namely assimilations almost exclusively occur with [coronal] segments that assimilate to either [labial] or [dorsal] place of articulation, but not the other way around.

The model can also explain the finding that in general, RTs are slower for coronals than for labials. During recognition, features are extracted from the speech signal and matched directly onto lexical representations. However, for [coronal] sounds, there is no feature in the lexical representation that can be matched. Therefore, a nonmismatch condition arises. On the other hand, [labial] features can be matched onto a labial feature in a lexicon. Although there is no direct claim that nonmismatch conditions are always slower than matching conditions, the results are not unexpected from a theoretical point of view. However, there need to be more studies in order to exactly determine in how far there is a time advantage for matching versus nonmismatching conditions.

This was a first attempt to examine the perceptual consequences of such assimilations in conversational speech and it appears that if a coronal consonant is assimilated to the following consonant, it is perceived as such even when presented without any context. Further studies are necessary to see if this pattern or results hold for other languages as well.

ACKNOWLEDGMENTS

This work was in part supported by the Deutsche Forschungsgesellschaft (SFB 471, SPP 1234, and the Leibniz Preis awarded to A.L.) and the Ministry of Science, Research, and the Arts of Baden-Württemberg. We also wish to thank our anonymous reviewers for their very useful comments and suggestions.

APPENDIX A: ALL PRONUNCIATION VARIATIONS OF *einverstanden* IN THE CORPUS IN IPA TRANSCRIPTION

	Phonetic transcription	Deviations from canonical transcription
	[ˈʔaɪnfɐˌʃtandən]	Canonical transcription, no deviations
i	[ˈa̰ɪ̰nfɐˌʃtan̰n]	1 segment deletion, 2 glottalizations
ii	[ˈa̰ɪ̰nfɐˌʃtanʔn]	1 segment deletion, 1 glottalization, 1 weakening
iii	[ˈa̰ɪ̰nfɐˌʃtan̰]	2 segment deletions, 2 glottalizations
iv	[ˈa̰ɪ̰nfɐˌʃtan]	3 segment deletions, 1 glottalization
v	[ˈa̰ɪ̰nfɐˌʃtandn]	1 segment deletion, 1 glottalization
vi	[ˈa̰ɪ̰nfɐˌʃtanhn]	1 segment deletion, 1 glottalization, 1 weakening
vii	[ˈaɪmfɐˌʃta̰n]	3 segment deletions, 1 glottalization, 1 assimilation
viii	[ˈʔa̰ɪ̰nfɐˌʃtan̰]	2 segment deletions, 2 glottalizations
ix	[ˈʔa̰ɪ̰nfɐˌʃtanʔn]	1 segment deletion, 1 glottalization, 1 weakening
x	[ˈʔa̰ɪ̰nfɐˌʃtan̰]	2 segment deletions, 2 glottalizations
xi	[ˈʔa̰ɪ̰nfɐˌʃtan̰n]	1 segment deletion, 2 glottalizations
xii	[ˈʔa̰ɪ̰nfɐˌʃtandn]	1 segment deletion, 1 glottalization
xiii	[ˈʔa̰ɪ̰ntfɐˌʃtanhn]	1 segment deletion, 1 glottalization, 1 weakening
xiv	[ˈʔaɪnfɐˌʃtanhn]	1 segment deletion, 1 weakening
xv	[ˈa̰ɪ̰nfɐˌʃtann]	2 segment deletions, 1 glottalization
xvi	[ˈa̰ɪ̰ntfɐˌʃtan̰]	2 segment deletions, 2 glottalizations, 1 insertion
xvii	[ˈa̰ɪ̰ntfɐˌʃtan̰n]	1 segment deletion, 2 glottalizations, 1 insertion
xviii	[ˈa̰ɪ̰nvɐˌʃtann]	2 segment deletions, 1 glottalization, 1 voicing
xix	[ˈʔaɪnfɐˌʃtanhn]	1 segment deletion, 1 weakening
xx	[ˈaɪnfɐˌʃanʔn]	2 segment deletions, 1 weakening
xxi	[ˈaɪnfɐˌʃta̰n]	3 segment deletions, 1 glottalization
xxii	[ˈaɪnfɐˌʃtan̰]	3 segment deletions, 1 glottalization
xxiii	[nfɐʃtan̰]	4 segment deletions, 1 glottalization

APPENDIX B: EXAMPLES OF CONTEXTS FROM WHICH [ɑm] AND [ɑn] STIMULI WERE EXTRACTED FOR EXPERIMENTS 1 and 2. FOR DETAILS SEE TABLE V

Kiel corpus Transcription	Example stimuli in orthography	Conditions
		Unassimilated-/m/
[ɑm]	…daram **am** e̲rsten…	Vowel context /m/-vowel
[ɑm]	…das denn **am** B̲esten	Labial context /m/-labial
[ɑm]	…wir **am** s̲echsten	Coronal context /m/-coronal
[ɑm]	…wir **am** günstigsten…	Dorsal context /m/-dorsal
		Phoneme-/n/
[ɑn]	…sieht das d**ann** a̲us…	Vowel context /n/-vowel
[ɑn]	**Dann** b̲rauchen wir…	Labial context /n/-labial
[ɑn]	Ist d**ann** d̲er…	Coronal context /n/-coronal
[ɑn]	…aber m**an** k̲ann…	Dorsal context /n/-dorsal
		Assimilated
[ɑm]	Und d**ann** b̲rauchen wir…	Labial context

[1]We ignored the "–h" symbol in the Kiel transcriptions since it has many phonetic correlates (e.g., aspiration, release) which are not relevant here. We translated the SAMPA transcription of the Kiel corpus into standard IPA transcription.

[2]We did not treat glottalization as instance of complete deletion, rather as some remnant of a severely reduced segment to keep the two processes apart.

[3]Neutralizations occur when speakers eliminate contrastive featural contrasts of segments in speech production. For instance, when they produce a segment such as /n/—underlyingly [coronal]—as a [labial] [m] due to a complete assimilation to the place of articulation of an upcoming [labial] segment, such as [b]. Complete means that the resulting [m] (underlyingly [n]) is not different from an underlying /m/ being produced as [m].

[4]Regressive assimilation occurs when in a sequence of two segments S_1 and S_2, S_1 assimilates in some feature(s) to S_2. Progressive assimilation occurs, when S_2 assimilates to S_1.

[5]Throughout this paper, we use the following convention for the description of letters and sounds. The sign ⟨ ⟩ is used to refer to orthography, [ ] indicates phonetic transcription, / / is used for underlying segments, and { } encloses morphemes.

[6]Along with Wiese (1996), Benware (1986) sees the "phonological word" as the only domain where regressive place assimilation can occur. He cites Kallmeyer (1981) for a case of regressive place assimilation in *kaputt gegangen* "has broken down," where the final /t/ of *kaputt* "ruined" is pronounced with a [k]. The phrase *kaputt gehen* consisting of two words is interpreted as a single phonological word in the sense that they form a very close unit, different from usual words in a phrase (Benware, 1986, p. 129).

[7]Since final devoicing affects all places of articulation, we did not differentiate between voiced versus voiceless segments.

[8]If there is a deletion and no assimilation on the preceding segment, it is not clear, whether the deleted segment itself was assimilated. If the preceding segment assimilates, it is not clear, whether the deleted segment triggered the assimilation, or the first segment of the upcoming word.

[9]The features are based on Lahiri and Reetz (2002). Palatals are assumed to be [coronal], as in many phonological accounts (e.g., Lahiri and Evers, 1991; Clements and Hume, 1995; Kenstowicz, 1994; for a different view, see for example Hall, 2000). The segments [x, ç] are assumed to be underlyingly placeless since the place of articulation of the preceding vowel determines the place of articulation of the fricative—[coronal] after front vowels, [dorsal] after back vowels. For sake of simplicity, we refer to the underlying fricatives as /x/ or /ç/.

[10]The fricative [x] is the only [dorsal] consonant function words end with. Due to final devoicing, only voiceless obstruents occur in C_1 target position.

[11]Almost all the cases of /m/ assimilating to /n/ could also be analyzed as being a wrong case-marking, a phenomenon that is well known for many German speakers (Bayer and Brandner, 2004; Schiering, 2005); *den* "the-*.accusative*" instead of *dem* "the.*dative*," etc. However, here we treated them as any other case of assimilation.

[12]We use the term congruent for responses where the transcription of the corpus was the same as the subjects' decision and incongruent for the opposite case.

[13]The analysis was carried out using SAS statistic software JMP, version 5.0.1.2.

[14]Since we were interested in the influence of the response on the reaction time, the responses are treated as a factor.

[15]The residual maximum likelihood (REML) estimation does not substitute missing values with estimated means and does not need synthetic denominators; rather the individual factors are tested against the whole model. This method is more conservative than the traditional expected mean squares estimation. Not significant results did not reach the 5% level.

[16]Subjects had only two possible response buttons, i.e., [n] or [m] to choose from in experiment 1. As can be seen, especially unassimilated-/n/ items in labial context produced a high amount of incongruent responses. This is arguably due to coarticulatory cues. For items in dorsal context, one could also expect coarticulatory cues influencing subjects' responses. However, it is not clear, how subjects would react in this situation, since there was no possibility to indicate "something else." In order to examine the nature of incongruent responses further, we opted for a free transcription task, where subjects could write what they heard without being restricted to two responses, in fact without being restricted to a nasal response at all.

Bayer, J., and Brandner, E. (**2004**). "Klitisiertes *zu* im Bairischen and Alemannischen," (Cliticized to in Bavarian and Alemannic) in *Morphologie und Syntax deutscher Dialekte und historische Dialektologie des Deutschen*, edited by F. Patocka and P. Wiesinger (Praesens Edition, Wien), pp. 160–188.

Benware, W. A. (**1986**). *Phonetics and Phonology of Modern German* (Georgetown University Press, Washington, DC).

Bybee, J. (**2007**). *Frequency of Use and the Organization of Language* (Oxford University Press, Oxford).

Clements, G. N. (**2001**). "Representational economy in constraint-based phonology," in *Distinctive Feature Theory*, edited by T. A. Hall (Mouton de Gruyter, Berlin), pp. 71–146.

Clements, G. N., and Hume, E. V. (**1995**). “The internal organization of speech sounds,” in *The Handbook of Phonological Theory*, edited by J. A. Goldsmith (Blackwell, Cambridge, MA).

Dilley, L. C., and Pitt, M. A. (**2007**). “A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition,” J. Acoust. Soc. Am. **122**, 2340–2353.

Dressler, W., Fasching, P., Chromec, E., Wintersberger, W., Leodolter, R., Stark, H., Groll, G., Reinhart, J., and Pohl, H. D. (**1972**). “Phonological fast speech rules in colloquial Viennese German,” Wiener Linguistische Gazette **1**, 1–30.

Ghini, M. (**2001**). *Asymmetries in the Phonology of Miogliola* (Mouton de Gruyter, Berlin).

Gow, D. W., Jr. (**2002**). “Does English coronal place assimilation create lexical ambiguity?,” J. Exp. Psychol. Hum. Percept. Perform. **28**, 163–179.

Hall, T. A. (**1992**). *Syllable Structure and Syllable Related Processes in German* (Max Niemeyer Verlag, Tübingen).

Hall, T. A. (**1999**). “Phonotactics and the prosodic structure of German function words,” in *Studies on the Phonological Word*, edited by T. A. Hall and U. Kleinhenz (John Benjamins, Amsterdam), pp. 99–131.

Hall, T. A. (**2000**). *Phonologie—eine Einführung* (Phonology—An introduction) (Walter de Gruyter, Berlin).

IPDS—Institut für Phonetik und digitale Sprachverarbeitung (**1994**). *The Kiel Corpus of Spontaneous Speech, CD-ROM I-III* (IPDS, Kiel).

Kabak, B., and Schiering, R. (**2006**). “The phonology and morphology of function word contractions in German,” J. Comparative Germanic Linguistics **9**, 53–99.

Kaisse, E. M. (**1985**). *Connected Speech: The Interaction of Syntax and Phonology* (Academic, Orlando).

Kallmeyer, W. (**1981**). “Aushandlung und Bedeutungskonstitution,” (Negotiation and constitution of meaning) in *Dialogforschung—Jahrbuch des Instituts für Deutsche Sprache, 1980*, edited by P. Schröder and H. Steger (Pädagogischer Verlag Schwann, Mannheim), pp. 89–127.

Kenstowicz, M. (**1994**). *Phonology in Generative Grammar* (Blackwell, Oxford).

Kohler, K. J. (**1990**). “Segmental reduction in connected speech in German: Phonological facts and phonetic explanations,” in *Speech Production and Speech Modelling*, edited by W. J. Hardcastle and A. Marchal (Kluwer Academic, Dordrecht), pp. 69–92.

Kohler, K. J. (**1995**). *Einführung in die Phonetik des Deutschen*, 2nd ed. (Introduction into the phonetics of German) (Erich Schmidt Verlag, Berlin).

Kohler, K. J., Pätzold, M., and Simpson, A. P. (Eds.) (**1995**). “From scenario to segment—the controlled elicitation, transcription, segmentation and labelling of spontaneous speech,” *AIPUK* (IPDS, Kiel), Vol. **29**.

Lahiri, A., and Evers, V. (**1991**). “Palatalisation and coronality,” in *The special Status of Coronals: Internal and External Evidence*, edited by C. Paradis and J.-F. Prunet (Academic, San Diego), pp. 79–100.

Lahiri, A., and Marslen-Wilson, W. D. (**1991**). “The mental representation of lexical form: A phonological approach to the recognition lexicon,” Cognition **38**, 245–294.

Lahiri, A., and Reetz, H. (**2002**). “Underspecified recognition,” in *Laboratory Phonology 7*, edited by C. Gussenhoven and N. Warner (Mouton de Gruyter, Berlin), pp. 637–675.

Local, J. (**2003**). “Variable domains and variable relevance: interpreting phonetic exponents,” J. Phonetics **31**, 321–339.

Manuel, S. Y. (**1995**). “Speakers nasalize /ð/ after /n/, but listeners still hear /ð/,” J. Phonetics **23**, 453–476.

Nespor, M., and Vogel, I. (**1986**). *Prosodic phonology* (Foris Publications, Dordrecht/Riverton).

Nolan, F. (**1992**). “The descriptive role of segments,” in *Papers in Laboratory Phonology II—Gesture, Segment, Prosody*, edited by G. J. Docherty and R. Ladd (Cambridge University Press, Cambridge), pp. 261–280.

OED (**1989**). *Oxford English Dictionary*, 2nd ed. (Oxford University Press, Oxford).

Ogden, R. A. (**1999**). “A declarative account of strong and weak auxiliaries in English,” Phonology **16**, 55–92.

Paradis, C., and Prunet, J.-F. (Eds.) (**1991**). *Phonetics and Phonology 2: The Special Status of Coronals: Internal and External Evidence* (Academic, San Diego).

Phillips, B. S. (**2001**). “Lexical diffusion, lexical frequency, and lexical analysis,” in *Frequency and the Emergence of Linguistic Structure*, edited by J. Bybee and P. J. Hopper (John Benjamins, Amsterdam), pp. 123–136.

Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., and Fosler-Lussier, E. (**2006**). “Buckeye Corpus of conversational speech,” (1st release), Department of Psychology, Ohio State University: Columbus, OH.

Reetz, H., and Kleinmann, A. (**2003**). “Multi-subject hardware for experiment control and precise reaction time measurement,” in Proceedings ICPhS Barcelona, pp. 1489–1492.

Scharinger, M. (**2006**). “The representation of vocalic features in vowel alternations. Phonological, morphological and computational aspects,” Ph.D. thesis, Konstanz, Konstanz Online Publication System, http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-24341 (Last accessed on 1/7/2008).

Schiering, R. (**2005**). “Flektierte Präpositionen im Deutschen? Neue Evidenz aus dem Ruhrgebiet,” (Inflected prepositions in German? New evidence from the Ruhr area), Zeitschrift für Dialektologie und Linguistik **72**, 52–79.

Selkirk, E. O. (**1984**). *Phonology and Syntax: The Relation Between Sound and Structure* (The MIT Press, Cambridge, MA).

Snoeren, N. D., Hallé, P. A., and Segui, J. (**2006**). “A voice for the voiceless: Production and perception of assimilated stops in French,” J. Phonetics **34**, 241–268.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (The MIT Press, Cambridge, MA).

Vater, H. (Ed.). (**1979**). *Phonologische Probleme des Deutschen* (Phonological problems of German) (Gunter Narr Verlag, Tübingen).

Wheeldon, L., and Waksler, R. (**2004**). “Phonological underspecification and mapping mechanisms in the speech recognition lexicon,” Brain Lang. **90**, 401–412.

Wiese, R. (**1996**). *The Phonology of German* (Oxford University Press, Oxford).

Wurzel, W. U. (**1970**). *Studien zur Deutschen Lautstruktur* (Studies on the German sound structure) (Akademie Verlag, Berlin).