Phonologische Identität und ihre Vermeidung

Einsichten aus Grammatik und Sprachverarbeitung

Habilitationsschrift

vorgelegt von Dr. Gerrit Kentner

Fachbereich 10 Neuere Philologien
Goethe-Universität Frankfurt

Dezember 2017

Danksagung

Zur Entstehung dieser Arbeit haben viele in vielfältiger Weise beigetragen. Mein ganz besonderer Dank gilt Caroline Féry, die meinen akademischen Werdegang ermöglicht und mit großem Vertrauen begleitet hat. Viele der in dieser Arbeit behandelten Themen sind von ihr inspiriert worden, stets habe ich von ihr Ermutigung erfahren und Freiraum für die Entwicklung eigener Ideen erhalten.

Dankbar bin ich auch für das anregende Umfeld der Frankfurter Linguistik und für die KollegInnen, die mich in den Jahren auf meinem Weg begleitet haben: Sascha Alexeyenko, Markus Bader, Leah Bauke, Julia Biskupek, Emilia Ellsiepen, Martin Elsig, Ingo Feldhausen, Eric Fuss, Pia Gerhard, Katharina Hartmann, Elke Höhe-Kupfer, Jan Köpping, Vasiliki Koukoulioti, Cécile Meyer, Anja Müller, Johannes Mursell, Sara Myrberg, Andreas Pankau, Cecilia Poletto, Yvonne Portele, Louise Probst, Henning Reetz, Esther Rinke, Petra Schulz, Matthias Schulze-Bünte, Manfred Sailer, Emanuela Sanfelici, Pete Smith, Frank Sode, Anke Stakemann, Thomas Strobel, Ewa Trutkowski, Dina Voloshina, Gert Webelhut, Helmut Weiß und Ede Zimmermann – sie alle haben zu dem überaus angenehmen Arbeitsklima in Frankfurt beigetragen. Hervorheben möchte ich hier die Gefährten aus der Phonologie und meine Büronachbarn, die sowohl für konstruktiv-kritische Diskussionen als auch für fröhliche Stunden der Ablenkung zur Verfügung standen: Caroline Féry, Isabelle Franz, Fatima Hamlaoui, Tim Hirschberg, Shin Ishihara, Beata Moskal, Sara Myrberg und Birgit Nutz – ich schätze mich glücklich, diese Menschen meine KollegInnen nennen zu dürfen.

Meiner Ko-Autorin Isabelle Franz und meinem Ko-Autor Shravan Vasishth danke ich für die fruchtbare Zusammenarbeit. Der Deutschen Forschungsgemeinschaft verdanke ich die Mittel für das Projekt "Prosodische Syntax des Deutschen" – einige Projektergebisse sind Teil dieser Habilitationsschrift. Christian Dück, Melanie Lüttin, Johannes Messerschmidt, Alex Pösch, Marc Schwab und Ani Sinaris haben als studentische Hilfskräfte zum Entstehen dieser Arbeit beigetragen. Ich danke ihnen und den vielen Versuchspersonen.

Jenny, Tilo, Olaf und Anne, ihr seid das große Glück in meinem Leben, Euch sei diese Arbeit gewidmet.

Inhaltsverzeichnis

Many songs have been written - and this is one of them.
Mo' Horinzons

1 Einleitung

Das Sprechen einer Sprache bedeutet, von "endlichen Mitteln unendlichen Gebrauch" zu machen (W. von Humboldt), man kommt daher als Sprachnutzer um die Wiederholung sprachlicher Einheiten (z.B. Wörter, Silben, Laute und die sie jeweils konstituierenden Merkmale) nicht umhin. Die Ausdrucksfähigkeit einer Sprache und die Grenzen ihrer Verarbeitbarkeit setzen allerdings voraus, dass die Wiederholung einen gewissen Schwellenwert nicht überschreitet. Die Verkettung immer gleicher Einheiten innerhalb einer Äußerung führt nämlich zu (zumindest formseitiger) Redundanz; im Extremfall wird eine Äußerung, die immer gleiche Einheiten verkettet, gar nicht oder nur nach einiger analytischer Anstrengung als sinnvoll wahrgenommen, selbst wenn sie einen grammatischen Satz darstellt (1).

(1) a. *Wenn Fliegen hinter Fliegen fliegen, fliegen Fliegen Fliegen nach.*
b. *Wer "brauchen" ohne "zu" gebraucht braucht "brauchen" gar nicht zu gebrauchen.*

Ein zu hoher Grad von Wiederholung kann auch artikulatorische Schwierigkeiten nach sich ziehen, wie die Zungebrecher in (2) zeigen.

(2) a. *Blaukraut bleibt Blaukraut und Brautkleid bleibt Brautkleid.*
b. *Der Kaplan klebt Pappplakate.*

Sätze dieser Art werden in aller Regel nicht spontan gebildet, ihre wörtliche Bedeutung steht in der kommunikativen Situation, in der sie geäußert werden, nicht im Vordergrund, sondern ihre poetische Funktion. Solche Sätze werden praktisch nur zitiert und zwar meist, um komische Effekte zu erzielen.

Das effektive Vermitteln von Botschaften erfordert im Normalfall, dass sprachliche Einheiten alternieren. Das Alternanzgebot gilt für die Laut- oder Segmentebene ebenso wie für höhere phonologische Domänen, wie z.B. die Silbe. Optimale Lautfolgen wechseln zwischen Konsonanten und Vokalen. Es gibt wohl keine Sprache, die innerhalb eines Wortes grundsätzlich alle Konsonanten vor alle Vokale reiht oder umgekehrt. Ähnliches kann für Sprachen, die Betonung nutzen, über betonte und unbetonte Silben gesagt werden.

Interessanterweise gilt das Alternanzgebot auch jenseits der Phonologie, z.B. auf der morpho-syntaktischen Ebene: normalerweise folgen im Satz Funktions- und Inhaltswörter (oder -morpheme) in annähernd regelmäßigem Wechsel aufeinander. Keine Grammatik verlangt, dass im Satz notwendigerweise alle Funktionswörter allen Inhaltswörtern vorangehen bzw. folgen. Hier sind offenbar domänenübergreifende Strukturprinzipien am Werk, die über die Verknüpfung asymmetrischer Einheiten - nämlich Kopf und Dependent - eine hierarchische Strukturierung des Sprachsignals erzeugen; dies hat im Regelfall Alternanz zur Folge. In der segmentalen Phonologie betrifft diese Asymmetrie das Verhältnis von Vokalen und Konsonanten; in der metrischen Phonologie das Verhältnis von betonten Kopfsilben und unbetonten Silben. In der Syntax u.a. das Verhältnis von Funktions- und Inhaltsmorphemen.

Auch auf der Diskursebene ist Identitätsvermeidung zu beobachten: Gordon et al. (1993) berichten von Leseexperimenten, die zeigen, dass der wiederholte Verweis auf einen Referenten zu erschwerter Textverarbeitung führt, wenn der Referent jeweils mit demselben Namen belegt wird – Gordon und Kollegen prägen dafür den Begriff *repeated name penalty*; bevorzugt wird dagegen nach erfolgter Nennung eines Namens oder einer Kennzeichnung die pronominale Referenz oder Alternanz zwischen namentlicher Nennung, Kennzeichnung und pronominaler Referenz.

In der phonologischen Literatur ist die Identitätsvermeidung unter dem Begriff "Obligatory contour principle", kurz OCP (McCarthy, 1986, u.v.m.) zu einem grammatischen Prinzip (3) erhoben worden.

(3) Obligatorisches Konturprinzip: Adjazente identische Elemente in der phonologischen Repräsentation sind markiert.

Die Formulierung ist dabei weitgehend offen gehalten, so dass das OCP sowohl die Adjazenz identischer Segmente (McCarthy, 1986; Schein and Steriade, 1986, und andere) und Autosegmente, wie Wörter (Golston, 1995) und Silben (Brugmann, 1917; Yip, 1988), Silbenkonstituenten (Plag, 1998), und Töne (Leben, 1973) verhindert. In der metrischen Phonologie gilt insbesondere das direkte Aufeinandertreffen betonter oder akzentuierter Silben (*stress clash*) und die Abfolge mehrerer unbetonter Silben (*stress lapse*) als markiert (Selkirk, 1984; Hayes, 1995; Schlüter, 2005).

Interessanterweise ist das OCP für höhere prosodische Domänen kaum diskutiert. Stattdessen gibt es jenseits der Silbe deutliche Evidenz für Selbstähnlichkeit und parallel aufgebaute Strukturen, Phänomene, die unter anderem auch aus der Musik bekannt sind (Pareyon, 2011). Wesentliche Teile der hier vorliegenden Arbeit sind mit Selbstähnlichkeit und phonologischer Identität in höheren prosodischen Domänen befasst (s. Abschnitte 3 und 4); aber es gibt auch auf segmentaler Ebene Tendenzen zu phonologischer Identität, die dem OCP-Prinzip offenbar zuwiderlaufen. So sind z.B. in der Kindersprache Konsonantenharmonie und Reduplikation recht häufig beobachtbar (Dressler et al., 2005; Fikkert and Levelt, 2008; Pater and Werle, 2003).

Die folgenden Abschnitte stellen für das Deutsche einen Überblick über Phänomene dar, die entweder phonologische Identität benachbarter Einheiten hervorrufen oder ihre Vermeidung zum Ziel haben. Zunächst betrachten wir in Abschnitt 2 die phonologische Identität und Identitätsvermeidung auf der segmentalen Ebene. Abschnitt 3 gibt einen Überblick über phonologische Identität(svermeidung) auf der suprasegmentalen Ebene. Der Vergleich dieser Ebenen zeigt, dass segmentale Identität generell vermieden wird, suprasegmentale Einheiten dagegen zu identischem Bau und Selbstähnlichkeit tendieren. Die in Abschnitt 2 und 3 vorgestellten Phänomene sind ohne Anspruch auf Vollständigkeit zusammengestellt, ich hoffe gleichwohl, eine repräsentative Auswahl getroffen zu haben, die Generalisierungen erlaubt. Es bleibt allerdings an dieser Stelle bei recht groben Generalisierungen, eine formale Analyse wäre bei anderer Gelegenheit noch zu leisten. In Abschnitt 4 gebe ich einen Überblick über die Schriften dieser kumulativen Habilitationsschrift.

2 Identität und Identitätsvermeidung auf segmentaler Ebene

Zu den identitätsstiftenden Phänomenen gehören auf der Segmentebene Assimilation und Silbendopplung; Degeminierung, Vermeidung von Konsonantenwiederholung und Haplologie dagegen wirken dissimilatorisch. Diese Prozesse werden im Folgenden näher vorgestellt und in Abschnitt 2.8 im Überblick diskutiert.

2.1 Degeminierung

Der Effekt des OCP zeigt sich in unterschiedlichen Strategien, die allesamt die Vermeidung multipler Strukturanwendung zum Ziel haben. Auf der Segmentebene ist zum einen die Vermeidung von Geminaten zu nennen. Im Deutschen sind Geminaten in nur wenigen Kontexten lizensiert, weshalb Sprecher entweder degeminieren (Hans' Pullover [hans pʊloːvɐ] ∼ *[hansː pʊloːvɐ]), Material einfügen (*Hansens Pullover*) oder eine Paraphrase bilden, die Geminierung gar nicht erst hervorruft (*der Pullover von Hans*). Nicht alle diese Strategien sind gleichermaßen zulässig. In (4-c) ist weder die Geminate noch Degeminierung lizensiert, Epenthese (*die Fussens*) ist möglich, scheint aber gegenüber der Paraphrase *Familie Fuss* markiert. (4-d) legt nah, dass Affigierung mit *-s* möglich ist, wenn das auslautende [s] im Namen nicht der einzige Kodakonsonant ist. In diesem Fall wird - wie beim Genitiv von *Hans* - obligatorisch degeminiert (s. Wiese, 1996, für weitere Fälle, in denen obligatorisch degeminiert wird).

(4) Familienname - Kollektivum
- a. *Müller - die Müllers*
- b. *Schmidt - die Schmidts*
- c. *Fuss - *die Fuss ∼ die Fussens ∼ Familie Fuss*
- d. *Jacobs - die Jacobs ∼ die Jacobsens ∼ Familie Jacobs*

Obwohl Geminaten stammintern und – wie eben gezeigt – im Kontext des Genitiv- oder Kollektivsuffix' *-s* im Deutschen nicht lizensiert sind, kommen Sie doch im Sprachstrom vor. Insbesondere die nasalen Flexionssuffixe */-n, -m/* neigen in normaler Rede dazu, mit stammfinalen Nasalen eine Geminate zu bilden (5), (6). Wenn mehrere gleichlautende Suffixe geminiert werden (5-b), geht dies mit Haplologie einher (s. Abschnitt 2.5). Allerdings ist die Geminierung dort, wo sie auftreten kann, nie obligatorisch: sie kann je nach Kontext und sprechtempoabhängig entweder durch schwa-Epenthese aufgebrochen werden oder durch Degeminierung vereinfacht werden (Wurzel, 1970).

(5) a. *beginnen*
[bəgɪnən] ∼ [bəgɪnː]
b. *begonnenen*
[bəgɔnənən] ∼ [bəgɔnːən]

(6) *(aus) armem (Hause)* [ʔaɐ̯məm] ∼ [ʔaɐ̯mː]

2.2 Assimilation

Die Tatsache, dass die Segmente im Sprachstrom nicht als deutlich voneinander abgegrenzte Einheiten verkettet werden, sondern die Artikulationsgesten benachbarter Laute sich zeitlich überlappen, sie also ko-artikuliert werden (Hardcastle and Hewlett, 2006, für einen Überblick), führt regelmäßig zu Assimilationen. Die Überlappung artikulatorischer Gesten betrifft in erster Linie direkt benachbarte Laute, kann aber, insbesondere im Fall der Vokalharmonie (die für das Deutsche synchron nicht relevant ist), auch über Konsonanten hinweg applizieren. Neben optionalen Assimilationen, wie sie z.B. sprechtempoabhängig über Grenzen phonologischer Wörter hinweg stattfindet (z.B. die Labialisierung des Nasals in *Weinbau* [vaɪ̯mbaʊ̯]), ist die regressive Nasalassimilation vor labialen und dorsalen Plosiven innerhalb prosodischer Füße obligatorisch (7-a), (7-b). Nur eine kleine Klasse von Wörtern zeigt in dieser Konstellation keine Assimilation (7-c); diese synchron nicht weiter analysierbaren Wörter weisen allerdings diachron eine Morphemgrenze zwischen dem für das Merkmal [labial] spezifizierten Nasal und dem Plosiv auf.

(7) a. *Kumpel* [kʊmpəl, *kʊnpəl, *kʊŋpəl]; *Wampe* [vampə, *vanpə, *vaŋpə]
b. *Bank* [baŋk, *bamk, *bank]; *Mango* [maŋgo, *mamgo, *mango]
c. *Imker* [ɪmkɐ] (zurückzuführen auf *Imme* +Suffix *-ker*);
Lemgo [lɛmgo] (wohl auf alte Gaubezeichnung *Limgauwe* zurückzuführen; siehe Köbler, 2007)

Im Standarddeutschen ebenfalls obligatorisch ist die regressive Assimilation des stimmlosen dorsalen Frikativs (8), der tautomorphemisch das Ortsmerkmal [+ hinten] eines direkt vorangehenden Vokals übernimmt (Kloeke, 1982; Macfarland and Pierrehumbert, 1991; Merchant, 1996; Noske, 1997, und andere).

(8) a. *Licht* [lɪçt, *lɪxt]; *weich* [vaɪ̯ç, *vaɪ̯x]
b. *Bach* [bax, *baç]; *Buche* [buːxe, *buːçə]

Synchron führt Assimilation innerhalb von Stämmen nicht zu kompletter phonologischer Identität und damit zu Geminierung der betroffenen Segmente (*Amnäsie* [amnesiː] ∼ [*anːesiː] ∼ [*amːesiː]). An der Grenze von Stamm und Affix oder über Stämme hinweg ist dies jedoch möglich (9), allerdings wie im Fall der Geminaten allgemein (Abschnitt 2.1) nie obligatorisch.

(9) a. *singen* [zɪŋən] ∼ [zɪŋː]
b. *Bannmeile* [banmaɪ̯lə] ∼ [bamːaɪ̯lə]

2.3 Vermeidung von Konsonantenwiederholung in einfachen Wurzeln

Auf der Eben der morphologischen Wurzeln ist eine Vermeidung von C_iVC_i-Strukturen, in denen ein Konsonantphonem wiederholt auftritt, beobachtbar. Dies scheint eine sprachübergreifende Tendenz zu sein, Pozdniakov and Segerer (2007) sprechen von einer statistischen Universalie (siehe auch Mayer et al., 2010). Beispielsweise ist für das Arabische bekannt, dass Wurzeln mit dieser Struktur grundsätzlich vermieden werden (Frisch et al., 2004; Pierrehumbert, 1993). Für das Deutsche zeigen Domahs et al. (2009), dass bestimmte Wurzeln mit Mehrfachvorkommen von Konsonanten nicht nur nicht belegt, sondern ungrammatisch sind. Das gilt insbesondere für Formen mit Konsonanten, denen im Ansatz ein weiterer Obstruent vorangeht [*ʃpɔp, *skiːk, *knan, *pleːl] – von diesem Verbot ausgenommen sind Formen mit koronalem Obstruenten (*Stadt, statt, Stätte*). Auf der Oberfläche sind die identischen Konsonanten aufgrund des eingeschobenen Vokals zwar nicht direkt adjazent. Die Autosegmentale Phonologie geht allerdings davon aus, dass Konsonanten und Vokale auf jeweils eigenen Segmentschichten repräsentiert sind (McCarthy, 1981, und viele andere). Auf der Konsonantenschicht sind die Konsonanten demnach adjazent, ungeachtet der auf der Oberflächenebene eingestreuten Vokale.[1] Die Behauptung, dass C_iVC_i-Strukturen im Deutschen markiert sind und vermieden werden, stützt sich auf eine eingehende Untersuchung des deutschen Kernwortschatzes. Richard Wiese (Marburg) hat mir dazu freundlicherweise eine auf Ortmann (1993) basierende Liste der deutschen Kernmorpheme oder Wurzeln zur Verfügung gestellt, die computerlesbar und damit bequem zu untersuchen ist. Diese Liste enthält insgesamt 6510 Einträge meist selbständiger Morpheme, die aus insgesamt acht Wörterbüchern zusammengestellt wurde (leider ist mir nicht bekannt, welche Wörterbücher genau dieser Sammlung zugrundeliegen). Für jedes Kernmorphem ist jeweils die orthographische und die phonologischen Gestalt verzeichnet. Für jeden Eintrag ist außerdem festgehalten, in wie vielen Wörterbüchern das entsprechende Morphem aufgezeichnet ist. Da für Lexika gilt, dass sie eher frequente und standardnahe Wörter auflisten, kann diese Zahl (minimal 1, maximal 8) als Heuristik für Standardnähe und Gebräuchlichkeit des entsprechenden Worts herhalten.

[1] Pierrehumbert (1993) kommt ohne die Annahme von voneinander getrennten Konsonant- und Vokalschichten aus und stellt stattdessen das Kriterium der relativen Nähe und der phonologischen Similarität der betroffenen Segmente in den Mittelpunkt ihrer Analyse der C_iVC_i-Vermeidung.

Methode: Die Liste wurde nach folgendem Schema analysiert. Zunächst wurde die Suche auf phonologisch einfache Stämme begrenzt, d.h. auf Wörter, die entweder einsilbig sind oder auf schwa-Silbe auslauten. Wörter mit zwei Vollvokalen, wie z.B. *Kuckuck* blieben unberücksichtigt. Für jede der übrigen 6375 Wurzeln wurde ermittelt, ob in ihrer Form Konsonanten mehrfach vorkommen. Ausschlaggebend für die Suche war die Wiederholung von Konsonanten auf der phonologischen Ebene. Dies gilt für insgesamt 327, d.h. ca 5% der Fälle. Von diesen weist ungefähr ein Drittel (n=111) adjazente identische Segmente auf der Konsonantenschicht auf.

(10) Wiederholung von Konsonanten

a. nicht-adjazent auf der Konsonantenschicht,
komplexe C-Strukturen: *Kalk, Start, Murmel*

b. adjazent auf Konsonantenschicht,
komplexe C-Strukturen: *Pfeffer, frier, Pups*

c. adjazent auf Konsonantenschicht,
einfache C_iVC_i(schwa)-Strukturen: *Puppe, Muhme, keck, Bibel*

Die folgenden Tabellen zeigen die Anzahl der Einträge nach Vorkommenshäufigkeit in den Quelllexika. Dabei zeigt sich, dass die Zahl der Wörter, die in nur einem von acht Lexika verzeichnet sind, in der Gruppe der Wörter mit Konsonantenwiederholung stärker repräsentiert sind (28% gegenüber 21% bezogen auf die Gesamtzahl der Kernmorpheme); umgekehrt sind 12% aller Wurzeln in allen acht Quelllexika aufgezeichnet (11); in der Gruppe der Wörter mit Konsonantenwiederholung (12) sind es allerdings nur noch 6%.

Eine ähnliche Verteilung zeigt sich bei den reinen C_iVC_i(schwa)-Strukturen, in denen es nur einen Konsonanten gibt, der aber zweimal vorkommt. Dies sind insgesamt 74 Fälle, von denen 19 (also 26%) in nur einem der acht Lexika verzeichnet sind. Diese Struktur kann daher als relativ ungebräuchlich eingestuft werden.

(11)

Lexikon gesamt	n=6510
Lemmas mit 8 Einträgen	774 (12%)
Lemmas mit 7 Einträgen	711 (11%)
Lemmas mit 6 Einträgen	719 (11%)
Lemmas mit 5 Einträgen	806 (12%)
Lemmas mit 4 Einträgen	833 (13%)
Lemmas mit 3 Einträgen	587 (9%)
Lemmas mit 2 Einträgen	733 (11%)
Lemmas mit 1 Einträgen	1347 (21%)

(12)

Strukturen mit Konsonantenwiederholung	n=327
Lemmas mit 8 Einträgen	20 (6%)
Lemmas mit 7 Einträgen	36 (11%)
Lemmas mit 6 Einträgen	29 (9%)
Lemmas mit 5 Einträgen	38 (12%)
Lemmas mit 4 Einträgen	39 (12%)
Lemmas mit 3 Einträgen	38 (12%)
Lemmas mit 2 Einträgen	35(11%)
Lemmas mit 1 Einträgen	92 (28%)

(13)

Reine C_iVC_i-Struktur	n=74
Lemmas mit 8 Einträgen	5 (7%)
Lemmas mit 7 Einträgen	6 (8%)
Lemmas mit 6 Einträgen	8 (11%)
Lemmas mit 5 Einträgen	10 (14%)
Lemmas mit 4 Einträgen	11 (15%)
Lemmas mit 3 Einträgen	8 (11%)
Lemmas mit 2 Einträgen	7 (9%)
Lemmas mit 1 Einträgen	19 (26%)

Bei näherer Betrachtung der reinen C_iVC_i-Morpheme fällt auf, dass viele der Formen nicht im standardsprachlichen Gebrauch sind; ein großer Teil hat onomatopoetischen Charakter oder entstammt einem kindersprachlichen oder kindgerichteten Register. Einige Formen sind offenbar lexikalische Atavismen aus frühen Stadien des kindlichen Spracherwerbs, für die bekannt ist, dass innerhalb von Wörtern Artikulationsortsmerkmale nicht kontrastieren (Fikkert and Levelt, 2008).

Ein Blick in ein etymologisches Lexikon untermauert diese Annahme ("openthesaurus" über `https://www.dwds.de`): Wir können hier vier Klassen von Wörtern unterscheiden, nämlich lautmalerische bzw. lautnachahmende Bildungen (14), Lexikalisierungen von Lallworten aus der Kindersprache (15), Entlehnungen (16) und auf ältere indoeuropäische Wurzeln zurückgehende Formen (17). Daneben ist die Etymologie (und die Bedeutung) für eine Reihe von Wörtern (18) unklar. Auch ohne expliziten Vergleich zu der Gruppe aller übrigen Wurzeln, die nicht die C_iVC_i-Struktur haben, kann man aufgrund der hohen Zahl von lautmalerischen und kindersprachlichen Bildungen davon ausgehen, dass die C_iVC_i-Wurzeln ein marginales Stratum des Lexikons repräsentieren. Dies kann als weiterer Beleg für Podzniakov & Segerers These gelten, wonach C_iVC_i-Strukturen universell vermieden werden.

(14) lautnachahmende Bildungen und sog. Schallworte
bubbern, dudeln, kakeln, pupen, Küken/Kücken, piepen/Pipe, tattern, tuten, geigen, sausen, säuseln, lallen

(15) auf Lallworte der Kindersprache zurückgehend
Bube, Pappe, pappern, Pappel, babben, Babe, Titte, Tüte, Tutte, Tütte(l), Tittel, Kacke, Muhme, Nonne, lullen, Memme, mümmeln

(16) Entlehnungen
Bob/bobben (engl.), Gig (engl.), Titel (lat.), Puppe/Püpp (lat.), Kick (engl.), Bibel (griech.), rar (lat.), Pope (russ.), Kak (obsolet, zu engl. 'cake'), kokeln (zu engl. 'coke'), Papel (lat.)

(17) auf ältere indoeurop. Wurzeln zurückgehend
bibbern, beben, Biber, Gugel, Kocke (obsolet), tot/töten, nennen/nann-, nein, neun, Tat/Täter, Ruhr, rühr, nun, keck

(18) unklare Etymologie
gegen, Gegend, kiek, kuck, Mumm, Rohr/Röhre, Sesel

2.4 Silbendopplung

Neben ihrem Vorkommen in lautmalerischen und kindersprachlichen Lall-Wurzeln findet sich die C_iVC_i-Struktur charakteristischerweise in Kosenamen des Typs *Jojo* < *Johannes*, *Kiki* < *Kirsten*. Diese Bildungen können als reduplizierte Trunkierungen beschrieben werden. Die Ausgangsform des Eigennamens wird auf eine leichte Silbe trunktiert, die - wohl um der Struktur das phonologische Gewicht eines Wortes zu verleihen (Saba Kirchner, 2010) - redupliziert wird (19-a). Trunkierung und Dopplung führen zur optimalen prosodischen Struktur des trochäischen Zweisilbers, die segmentale Struktur entspricht aber der universal markierten C_iVC_i-Abfolge. Dieser morphophonologische Prozess scheint schwach produktiv zu sein und ist deutlich durch den segmentalen Kontext beschränkt: So können weder Affrikaten, komplexe Ansätze, Glottallaute noch r-Laute die C-Position besetzen (19-b)-(19-f).

(19)
a. *Johannes* → [jo] → [jojo]
b. *Zacharias* → *[tsa] → *[tsatsa]
c. *Britta* → *[bri] → *[bribri]
d. *Ina* → *[ʔi] → *[ʔiʔi]
e. *Hartmut* → *[ha] → *[haha]
f. *Robert* → *[ro] → *[roro]

Auch für den Vokal gelten Beschränkungen: Die Silbendopplung appliziert offenbar nicht mit Diphthongen (20-a) oder gerundeten Vorderzungenvokalen (20-b).

(20) a. *Meike* → *[maɪ̯] → *[maɪ̯maɪ̯]
b. *Lydia* → *[ly] → *[lyly]
Hoeneß → *[hø] → *[høhø]

Trotz ihrer (schwachen) Produktivität bleiben die durch Silbendopplung gebildeten Formen einem Randbereich der Morphologie verhaftet, der sich durch seinen expressiven Charakter auszeichnet – es handelt sich nämlich ausschließlich um Namen, meist Koseformen, die auf diese Weise gebildet werden. Möglicherweise steht die Expressivität dieser Bildung gerade im Zusammenhang mit ihrer segmentalen Struktur, die ja für lautmalerische und kindersprachliche Wurzeln charakteristisch ist.

2.5 Haplologie

Das Gegenstück zur Silbendopplung ist die Haplologie. Die Haplologie ist ein dissimilatorisches Verfahren zur Vermeidung von adjazentem identischen Material in komplexen Wörtern; sie führt zu Formen, deren syllabische Struktur um eine Silbe kürzer erscheint als nach Anschauung der zugrundeliegenden Formen erwartbar wäre. Im Deutschen betrifft die Haplologie in der Regel unbetonbare Silben, mitunter können aber auch unbetonte, betonbare Silben betroffen sein (21).

(21) Morphophonologie ∼ Morphonologie

Regelhaft sind die im Stamm auf *-er* auslautenden *-er*-affigierten Nomina von Haplologie betroffen, wenn sie um das Movierungssuffix *-in* erweitert werden.

(22) (Beispiele aus Plag (1998))
a. Zauber-er-in → Zauberin
b. Bewunder-er-in → Bewunderin
c. Haiger-er-in → Haigerin

Plag (1998) nimmt an, dass diese Formen phonologischen Beschränkungen über die Struktur von Nuklei benachbarter Silben unterliegen. In Plags Analyse verhindert die Beschränkung OCP$_{nucleus}$ das Aufeinanderfolgen von zwei schwa-Silben, wie in (23) gezeigt. Diese OCP-Beschränkung markiert auf der Vokalschicht adjazente identische Silbennuklei.

(23) Plags Analyse der Haplologie im Deutschen

/tsaʊbr/-/r/-/ɪn/	OCP$_{nucleus}$	*COMPLEX$_{onset}$	MAX	DEP
tsaʊbɐʁɪn			*	*
tsaʊbɐʁɐʁɪn	*!			**
tsaʊbʁɐʁɪn		*!		*

Allerdings schließt Plags OT-Analyse auch durchaus grammatische Strukturen aus, nämlich alle Formen mit aufeinanderfolgenden Silben mit identischem Nukleus. Eine Eingabeform wie /tsaʊbr/-/r/ sollte nach dieser Grammatik, um OCP$_{nucleus}$ zu genügen, auch ohne Movierungssuffix zu *[tsaʊbɐ] haplologiert werden. Neben Wörtern mit aufeinanderfolgenden schwa-Silben (*Zauberer*) wären auch Wörter wie *Banane* betroffen. Wie zu zeigen sein wird, erfasst eine Theorie der Haplologie, die neben der segmentalen Struktur auch die Betonungsstruktur berücksichtigt, die Daten besser. Plag allerdings schließt eine betonungsbasierte Grammatik der Haplologie explizit aus. Als Kronzeugin für seine rein segmentbasierte Analyse ruft Plag Formen wie *Treburerin* auf (Plag 1997: 208), bei der keine Haplologie appliziert, obwohl hier nach Plag wie in **Zaubererin* drei unbetonte Silben aufeinanderfolgen. Da die Vokalsegmente in *Treburerin* alternieren, so Plag, verletzt dieses Muster OCP$_{nucleus}$ nicht - entsprechend bleibt sie von Haplologie verschont. Das Beispiel will allerdings nicht recht überzeugen: der Ortsname *Trebur* und die darauf basierenden Demonyme *Treburer* und *Treburerin* haben meines Erachtens zwei betonte Silben, zumindest zwei betonbare Silben. Auch wenn der Wortakzent auf die Initialsilbe fällt, so müssen der zweiten Silbe aufgrund des Vollvokals und dem auslautenden /r/ doch mindestens zwei Morawerte zugeschrieben werden. Ich nehme also für den Ortsnamen und die Demonyme die Betonungsstruktur *Trébùr, Trébùrer, Trébùrerin* an. Eine betonungsbasierte Analyse der Haplologie, die mehr als zwei aufeinanderfolgende unbetonte Silben vermeidet, sagt also genauso wenig wie die segmentbasierte Analyse Haplologie in diesem Fall voraus.

Die betonungsbasierte Analyse gewinnt aber gegenüber Plags segmentbasierter angesichts der vielen grammatischen Formen mit zwei aufeinanderfolgenden schwa-Silben. Ortsnamen und die dazugehörigen Demonyme sind in diesem Zusammenhang wieder zur Veranschaulichung geeignet.

Mithilfe einer Beschränkung, die ein daktylisches Maximum vorsieht, sodass höchstens zwei unbetonte Silben auf eine betonte folgen, würde die Grammatikalität von *Zauberer* und die Ungrammatikalität von **Zaubererin* richtig vorhergesagt. Allein macht eine solche Beschränkung allerdings noch nicht die korrekte Vorhersage: Formen wie *Aachenerin, Gießenerin, Bingenerin* etc. belegen dies. Drei unbetonte Silben sind also am rechten Wortrand grundsätzlich zulässig. Wichtig scheint, wie von Plag (1998) bereits erkannt, dass die segmentale Struktur der unbetonten Silben ausreichend variiert. Anders als von Plag vorhergesagt, scheint aber nicht der Nukleus, sondern der Reim bzw. die Koda ausschlaggebend. Es lässt sich nämlich folgende Generalisierung aus den Daten ableiten:

(24) Zwei phonologisch (nahezu) identische Silbenreime/Kodas sind in direkt aufeinanderfolgenden, unakzentuierten Silben nur am Wortende lizensiert.

Wie oben gesehen, ist /-r-r/ → [-ɐ.ʁɐ] (*Zauberer*) am Wortende grammatisch, /-r-r-ɪn/ → *[-ɐ.ʁɐ.ʁɪn] dagegen führt zu Haplologie (vgl. *Zauberin*). Ähnliches gilt für auf Nasal auslautende unbetonte Silben: Ortsnamen wie *Bevensen* [beː.vən.zən] oder *Tübingen* [tyː.bɪŋən] werden trunkiert, wenn sie um das *-er*-Suffix erweitert werden: *Bevenser* ∼ **Bevensener; Tübinger* ∼ **Tübingener*. Das Quasi-Minimalpaar **Tübingener(in) – Bingener(in)* zeigt besonders anschaulich, dass die Position der Hauptbetonung ausschlaggebend ist: die beiden Wörtern gemeinsame zugrundeliegende Segmentabfolge /-ɪŋənəʁ(ɪn)/, die zwei auf der Konsonantschicht adjazente Nasale enthält, erscheint nur dann transparent an der Oberfläche, wenn die akzenttragende Silbe den schwa-Silben direkt vorangeht. Dies ist in *Bingener(in)* der Fall; ansonsten kommt es wie bei **Tübingenerin* ∼ *Tübingerin* zu Haplologie.

2.6 Vermeidung segmentaler Identität bei Reduplikation, Selbstkomposition und Paarformeln

Auch auf höheren morphophonologischen und morphosyntaktischen Ebenen ist die Wirkung des OCP auf der Segmentebene belegbar. Der Fall der Reduplikation im Deutschen zeigt, dass phonologische Identität adjazenter Füße mithilfe von Reim (25-a) oder Ablaut (25-a) vermieden werden. Für Details verweise ich auf die Ausführungen in Kentner (2017).

(25) a. *Schickimicki, doppelmoppel, Hasepase*
a'. **schickischicki, *doppeldoppel. *Hasehase*
b. *Krimskrams, Quitschquatsch, Mischmasch*
b'. **Kramskrams, *Quatschquatsch, *Mischmisch*

Eine ebenfalls scheinbar reduplizierende Struktur haben die sogenannten Selbstkomposita (26). Für diese werden zwei identische Stämme verknüpft. Morphologisch handelt es sich bei diesen Wörtern aber nicht um Reduplikation im engeren Sinne, sondern um Rektionskomposita, bei denen der relationale Kopf (das Zweitglied) eine Argumentstelle eröffnet, die durch ein gleichlautendes Erstglied besetzt wird. Auch hier wird direkte Adjazenz der identitischen Stämme verhindert. In diesem Fall allerdings nicht durch segmentale Änderung der Stämme mittels Reim oder Ablaut, sondern indem die Kompositionsgrenze grundsätzlich ein Fugen *-(e)s* enthält.

(26) *Kindeskind, Enkelsenkel, Freundesfreund, Helfershelfer*

Auf Phrasenebene finden wir schließlich die meist idiomatischen Paarformeln (27-a) (Müller, 1997), die formseitig immer "zweieiig" sind. Auch wenn die durch *und* verbundenen Worte quasi-synonym sind, sind denkbare Konstruktionen mit identischen Konjunkten (27-b) nicht zulässig oder zumindest nicht pragmatisch äquivalent zu den Formen in (27-a).

(27) a. *holtern und poltern, mit Sack und Pack, hegen und pflegen*
b. **poltern und poltern, *mit Sack und Sack, *pflegen und pflegen*

In diesen Phrasemen verhindert das koordinierende *und* die direkte Nachbarschaft der phonologisch korrespondierenden Konjunkte; es besteht also anscheinend keine Notwendigkeit für den Wechsel der Ansatzkonsonanten zur Vermeidung von Gleichklang adjazenter Einheiten. Allerdings ist es bei Beibehaltung des idiomatischen Charakters möglich, die Konjunkte auch ohne *und* als Teil einer Aufzählung zu verwenden (28), zumindest solange die Reihenfolge der Konjunkte fixiert bleibt.

(28) Belege aus dem Internet
a. *Klinikumzug mit Sack, Pack und Patienten*
b. *...echte Oldtimer, die wir [...] hegen, pflegen und bewahren.*
c. *...als die IT-Abteilung noch schalten, walten und horten konnte*

Das *und* ist daher wohl nicht notwendiger Bestandteil des idiomatischen Ausdrucks – ein Umstand, der die segmentale Alternanz zur Gleichklangvermeidung motiviert. Dagegen scheint es weniger gut möglich, die idiomatischen Phraseme anderweitig aufzubrechen (29) (die Grammatikalitätsurteile sind, weil rein introspektiv, mit Vorsicht zu genießen und bedürfen weiterer Absicherung). Die phonologisch korrespondierenden Elemente scheinen also lokal aneinander gebunden, und können höchstens durch *und* voneinander getrennt werden.

(29) a. **mit Sack und schwerem Pack*
b. ?*wir hegen die Beziehung und pflegen sie*
c. **schalten und autoritär walten*

In diesem Zusammenhang ist eine weitere Konstruktion von Interesse (30), bei der charakteristischerweise phonologisches Material wiederholt wird, im Unterschied zu den Paarformeln (27-a) allerdings ohne segmentale Alternanz. Diese Konstruktion drückt Sequenzialität aus und weist ähnliche Lokalitätsbedingungen auf wie die Paarformel (Müller, 2016).

(30) a. *Tag für Tag*
b. *Jahr um Jahr*
c. *Seit' an Seit'*
d. *Stein auf Stein*

Anders als beim *und* in den Paarformeln ist es hier in keinem Fall möglich, die Präposition zwischen den Zwillingsnomina zu löschen. Adjazenter Gleichklang wird hier also mittels intervenierendem Material verhindert, ähnlich wie bei den Selbstkomposita (26), die notwendigerweise ein Fugenelement aufweisen.

2.7 Gleichklangvermeidung als Merkmal des Lexikons

Allen bisher genannten Strukturen ist gemein, dass sie lexikalisiert sind oder zumindest lexikalisiert werden können; das gilt nicht nur für die wortwertigen Reduplikationen (25) und die Komposita (26), sondern auch für die phrasalen Konstruktionen (27-a), (30).

Ich behaupte nun, dass die Identitätsvermeidung gerade ein Merkmal des Lexikons ist. Ein Blick auf okkasionelle Konstruktionen (31), in denen Adjazenz phonologisch identischer Elemente beobachtbar ist, belegt dies eindrücklich. Bei diesen Konstruktionen handelt es sich nämlich um welche, die typischerweise nicht lexikalisiert werden.[2] Im Allgemeinen sind diese Konstruktionen nur in einem Diskurskontext lizensiert, in dem das wiederholte Element bereits vorerwähnt wurde oder zumindest salient ist und nur in diesem Kontext werden sie, um Kontrast zum vorerwähnten oder salienten Element auszudrücken, *ad hoc* gebildet.

(31) a. Kontrastive Fokusreduplikation (Finkbeiner, 2014; Freywald, 2015, und dortige Referenzen):
Nimmst Du Reis-Reis oder Basmatireis?
(*Reis-Reis*=prototypischer Reis, i. Ggs. zu Basmatireis o.ä.)

b. X-und-X-Konstruktion (Finkbeiner, 2012):
A: *Du bist aber dünn geworden!*
B: *Naja, dünn und dünn...* (präsupponiert zwei deutlich unterschiedliche, implizite Vergleichsstandards zum Adjektiv)

2.8 Zusammenfassung und Diskussion

Der Überblick über die identitätsstiftenden und identitätsvermeidenden Prozesse auf der segmentaler Ebene ergibt interessante Einsichten in die Struktur der deutschen Wortformen. Adjazenz von identischen Konsonanten wird offenbar grundsätzlich vermieden und ist nur in wohldefinierten Randbereichen zulässig. Wenn identische Konsonanten auf der Segmentschicht in der zugrundeliegenden Repräsentation direkt adjazent liegen, wird innerhalb von einfachen oder flektierten Wörtern häufig degeminiert oder schwa eingestreut. Geminaten sind häufiger über Wortgrenzen hinweg, z.B. an der Kompositionsgrenze beobachtbar (*kusssüchtig* [kʊsːʏçtɪç]). Assimilation wirkt zwar hinsichtlich phonologischer Merkmale identitätsstiftend, indem z.B. ein Nasal das Ortsmerkmal des benachbarten Konsonanten

[2] Ausnahmen bestätigen hier die Regel: Der Markenname *Film-Film* hat es wohl ins Lexikon geschafft.

übernimmt – sie führt aber stammintern nicht soweit, dass benachbarte Segmente komplett identisch würden.

Wurzeln mit identischen Konsonanten, die auf der Segmentschicht nicht direkt adjazent stehen, aber lediglich durch vokalische Segmente voneinander getrennt sind, gelten als markiert. Das Lexikon enthält durchaus solche Formen – sie gehören aber zum großen Teil dem expressiven Wortschatz an oder sind auf expressive und onomatopoetische Formen zurückzuführen. Wenn solche Wörter durch Trunkierung und Silbendopplung neu gebildet werden, dann handelt es sich um ebenfalls expressive Koseformen. Die markierte phonologische Form weist gewissermaßen auf den expressiven Charakter dieser Wörter hin. Im Übrigen gibt es bei Koseformen wie *Jojo* < *Johannes* einen deutlichen prosodischen Unterschied zwischen den segmental identischen Silben, von denen nämlich die erste betont und die zweite unbetont ist – diese Formen entsprechen also grundsätzlich der trochäischen Struktur, die Stämme im Deutschen typischerweise aufweisen (Eisenberg, 1991; Féry, 1997, und andere).

Abfolgen von unakzentuierten Silben mit identischem Reim sind nur in bestimmten derivierten Umgebungen toleriert. Dort, wo diese Abfolgen durch Derivation möglich und nötig sind (am Wortende: *Zauberer*), rufen sie zwar – anders als die oben besprochenen Koseformen – keine sekundäre expressiven Bedeutungskomponente hervor. Ihr beschränktes Vorkommen (sie folgen offenbar nicht wortintern vor einer weiteren unbetonten Silbe direkt aufeinander **Zaubererin*) bestätigt aber die Vermeidung phonologischer Identität innerhalb von Wörtern auch oberhalb der Segmentbene.

Soweit komplexere Formen (Reim- und Ablautreduplikation, Komposita, Phrasen) lexikalisiert werden, so gilt auch für sie, wie in 2.6 gezeigt, Vermeidung segmentaler Identität adjazenter Einheiten. Konstruktionen, für die phonologische Identität charakteristisch ist (31), werden *ad hoc* gebildet und nicht oder nur schwer ins Lexikon übernommen.

3 Phonologische Identität(svermeidung) auf suprasegmentaler Ebene

3.1 Gebot der rhythmischen Alternanz

Das OCP wird grundsätzlich offen formuliert und gilt nicht nur auf der segmentalen Ebene. Das rhythmische Alternanzgebot fordert, dass Adjazenz von Silben mit (nahezu) identischem Betonungsstatus unabhängig von ihrem segmentalem Gehalt vermieden werden – und zwar wortintern und über Wortgrenzen hinweg. Sowohl Abfolgen von unbetonten Silben (*stress lapse*), als auch Abfolgen von betonten Silben (*stress clash*) gelten hinsichtlich dieses Gebots als markiert.

Die Grammatik stellt unterschiedliche Verfahren zur Verfügung, die auf Alternanz von betont und unbetont hinwirken. Haplologie zur Vermeidung von *stress lapse* ist schon angesprochen worden.

Mitunter sind morphologische Prozesse blockiert, wenn z.B. das Ergebnis der morphologischen Operation eine Form mit zwei unbetonten Silben am Wortanfang hervorrufen würde (32-c) (s.a. Vogt, 2013).

(32) Nominalisierung mit Prä- bzw. Zirkumfix *Ge-(-e)*
  a. red → Gerede
     wink → Gewinke
  b. säusel → Gesäusel
     bibber → Gebibber
  c. studier → *Gestudiere
     vergess → *Gevergesse

Die Pluralmorphologie im Deutschen (33) sieht je nach prosodischer Struktur des Stamms silbische (a.) und nicht-silbische Pluralsuffixe (b.) vor, um die Abfolge betont-unbetont am Wortende sicherzustellen (Eisenberg, 1991; Wegener, 2004).

(33)

	Singular	Plural
a.	Béet	Béet-e
	Ménsch	Ménsch-en
	Elefánt	Elefánt-en
b.	Kábel	Kábel-0
	Blúme	Blúme-n
	Mandaríne	Mandaríne-n

Phonologische Verfahren zur Vermeidung von *stress clash* sind Betonungslöschung und -verschiebung, die wir z.B. bei der Komposition (34-a), aber auch jenseits des Wortes innerhalb phonologischer Phrasen (34-b) beobachten können (Kiparsky, 1966; Henrich, 2015).

(34) a. *Bauarbeiter*: /ˈbaʊ̯ + ˈarbaɪ̯tr/ → [ˈbaʊ̯ɐˌbaɪ̯tɐ]
b. *das Hémd + ánziehen → das Hémd anzíehen*

Darüber hinaus wird berichtet, dass Sprecher Periphrasen formulieren oder syntaktische Umstellungen vornehmen, um rhythmische Alternanz zu gewährleisten. Anstelle des vorangestellten Genitivs wie in *Páuls Búch* würde beispielsweise die Struktur mit Präpositionalphrase *das Búch von Pául* die Adjazenz betonter Silben vermeiden (Kentner, 2018). Anttila et al. (2010) und Shih et al. (2015) weisen für das Englische einen Einfluss der rhythmischen Umgebung auf die Wahl zwischen syntaktisch unterschiedlichen, synonymen Konstruktionen nach. Vogel et al. (2015) zeigen, dass Sprecher flexibel linearisierbare Pronominaladverbien im Mittelfeld im Sinne einer rhythmischen Optimierung einsetzen und relativ rhythmische Strukturen wie (35-a) gegenüber weniger rhythmischen Abfolgen (35-b) bevorzugen.

(35) a. *Da wóllte der Péter Tomáten drin kóchen.*
b. *Da wóllte der Péter drin Tomáten kóchen.*

3.2 Phonologischer Parallelismus

Ein mit dem Gebot der rhythmischen Alternanz konfligierendes Wohlgeformtheitsgesetz ist das Parallelismusgebot (auch als Balance-Gebot geläufig, s. Breen et al., 2011) (36).

(36) Parallelismusgebot: Jede prosodische Domäne enthält eine Folge gleichartig strukturierter, d.h. parallel aufgebauter Subdomänen.

Dieses Gebot läuft identitätsvermeidenden Gesetzen grundsätzlich zuwider, weil es prosodischen Gleichklang benachbarter phonologischer Einheiten fordert. Wiese and Speyer (2015) und Wiese (2016) zeigen, dass mitunter Abfolgen von zwei betonten Silben bevorzugt werden, um phonologischen Parallelismus zu fördern. Ihre Untersuchungen belegen, dass parallele Strukturen mit zwei Einsilbern (z.B. *wär gern*) oder zwei Zweisilbern (*wäre gerne*) gegenüber asymmetrischen Kombinationen von Ein- und Zweisilber (*wäre gern, wär gerne*) bevorzugt werden. Im Fall der Kombination von zwei Einsilbern würde also ein *stress clash* zugunsten des Parallelismus in Kauf genommen. Ich verweise auf die Ausführungen in Kentner (2015) und Kentner (2018), in denen ich das von Wiese and Speyer (2015) vorgeschlagene Parallelismusgebot und seine Wirkung auf der Wort- und Fußebene eingehend diskutiere.

Die Formulierung des Parallelismusgebots in (36) ist – ähnlich wie das OCP offen gewählt, sodass es auch höhere prosodische Ebenen berührt. Und tatsächlich wurde vorgeschlagen, dass phonologische Phrasen idealerweise genau zwei Wörter ungefähr gleichen phonologischen Gewichts enthalten (Myrberg, 2013). Nach Selkirk (2000) ergibt sich eine solche Präferenz aus dem Zusammenwirken von zwei Beschränkungen, nämlich BINMIN (eine phonologische Phrase enthält minimal zwei prosodische Wörter) und BINMAX (eine phonologische Phrase enthält maximal zwei prosodische Wörter). Auch für Intonationsphrasen wird behauptet, dass sie idealerweise symmetrisch aufgebaut sind, d.h. gleichschwere phonologische Phrasen enthalten (Ferreira, 2002; Ghini, 1993).

Das Parallelismus- oder Similaritätsgebot als Sprachstrukturprinzip ist für die Sprachverarbeitung relevant: Grosjean et al. (1979) zeigen in einem Experiment, dass Leser Pausen in ungefähr gleichen Abständen bevorzugen (s.a. Cooper and Paccia-Cooper, 1980; Gee and Grosjean, 1983). Schweitzer et al. (2011) belegen Ähnliches für die mündliche Sprachproduktion anhand eines Nachrichtenkorpus. In ihrer Untersuchung finden sie eine deutliche Korrelation der Länge einer phonologischen Phrase mit der Länge der folgenden Phrase. Sprecher richten also die Phrasierung des Textes so aus, dass die Phrasengrenzen in regelmäßigem Abstand auftreten. Auch in der Sprachrezeption ist Parallelismus wirksam: Frazier and Fodor (1978) und Fodor (1998) nehmen an, dass ein erster Schritt in der Satzverarbeitung beim Lesen darin besteht, das Eingabematerial in ungefähr gleich große *chunks* aufzuteilen, für die dann die syntaktische Struktur zunächst jeweils separat berechnet wird. Dieses Vorgehen, das als SAME-SIZE-SISTER-Prinzip bzw. als Prinzip der SAUSAGE MACHINE Eingang in die Literatur gefunden hat, kann wohl als eine prosodische Heuristik aufgefasst werden, denn das Kriterium der Phrasengröße stellt keine syntaktische oder semantische Kategorie dar.

3.3 Diskussion

Während wir auf der segmentalen Ebene vor allem identitätsvermeidende Prozesse und Phänomene beobachten können, scheint auf der suprasegmentalen Ebene eine Neigung zu prosodischer Identität vorzuherrschen. Dies gilt für die Struktur phonologischer Phrasen, aber auch für die Wortebene. Eine interessante Interaktion von segmentaler Identitätsvermeidung und prosodischem Parallelismus zeigt sich bei Reim- und Ablautreduplikation im Deutschen (Kentner, 2017). Redupliziert werden nämlich Einheiten, die den Status eines prosodischen Fußes haben, wobei dieser hinsichtlich seiner internen Struktur unterspezifiziert ist: Es können sowohl Einsilber (*misch* → *Mischmasch*), als auch Trochäen (*Hase* → *Hasepase*) redupliziert werden. Das Ergebnis ist bei segmentaler Variation (Reim bzw. Ablaut) aber immer ein prosodisch paralleler Aufbau: Wir finden nur Zwei- oder Viersilbige Reduplikationen, nie asymmetrische Dreisilber (**Hasepas, *Mischmasche*). Das Identitätsverbot bei Reduplikation kann also nicht auf prosodischer Ebene erfüllt werden, sondern ist auf die Segmentebene beschränkt.

Auf der Ebene des prosodischen Fußes stellt die asymmetrische Verknüpfung von betont und unbetont das verbreitetste Muster dar. Hier haben Parallelismus-Prinzipien wie Myrbergs EQUAL SISTERS-Beschränkung oder Fodors SAME SIZE SISTER-Gebot nicht die Wirkung, die ihnen auf höheren prosodischen Ebenen zugeschrieben werden. Dort, wo segmental identische Silben aufeinanderfolgen, wie z.B. bei den Koseformen *Jojo* < *Johannes* oder sog. Lallworten (wie *Mama, Pipi* etc.), sind diese zumindest prosodisch deutlich hinsichtlich ihres Betonungsstatus unterschieden.

Die Zusammenschau der hier besprochenen Phänomene legt den Schluss nahe, dass die Neigung zu phonologischer Identität und ihre Vermeidung kaum in direktem Konflikt stehen. Adjazenz identischer phonologischer Einheiten wird auf der Segmentebene vermieden, wenn die Einheiten Teil eines lexikalisierten Ausdrucks sind oder der Ausdruck potentiell lexikalisierbar ist. Dabei scheint es unwesentlich, welche Größe die segmental identischen Einheiten haben: Identitätsvermeidung ist sowohl für benachbarte Segmente (Degeminierung), Silben (Haplologie), einfache Wurzeln (Markiertheit von C_iVC_i-Strukturen), Füße (Reim- und Ablautreduplikation), als auch für phrasale Einheiten wie die Paarformeln vom Typ *hegen und pflegen* nachweisbar. Für Randbereiche des Lexikons (expressive Ausdrücke, Koseformen) gilt dieses Identitätsverbot allerdings nur eingeschränkt. Das Identitätsverbot ist offenbar im Fall von kontextuell gebundenen, *ad hoc* gebildeten Ausdrücken, die in der Regel nicht lexikalisiert werden, nicht wirksam, wie die Beispiele in (31) zeigen.

Jenseits und unabhängig von der Segmentebene werden selbstähnliche, parallel aufgebaute prosodische Strukturen bevorzugt: Längere Äußerungen bestehen aus ungefähr gleich großen Intonationsphrasen; Intonationsphrasen bestehen aus möglichst gleich großen phonologischen Phrasen, diese wiederum aus Verknüpfungen von (idealerweise zwei) gleichartig strukturierten Wörtern oder Füßen.

Die einzige Ebene, auf der es einen Konflikt zwischen Gleichklangvermeidung (OCP) und dem Parallelismusgebot gibt ist die Wort- oder Phrasenebene oberhalb der Fußebene: Werden im Sinne des Parallelismusgebots zwei Einsilber mit Betonung zu einem komplexen Wort (*/hoch/+/Haus/* → *Hochhaus*) oder zu einer Phrase (*/schön/+/hell/* → *schön hell*) kombiniert, kommt es auf Kosten des Prinzips der rhythmischen Alternanz zum *stress clash*. Der Konflikt wird dann über das Kulminativitätsprinzip (Hyman, 2006; Riad, 2012) gelöst; dieses besagt, dass Wörter oder Phrasen nur ein prosodisches Maximum haben. Im Deutschen bedeutet dies für Komposita in der Regel Prominenz des linken Glieds, und im Fall von Phrasen rechtsseitige Prominenz (Wiese, 2001). Bei der Verknüpfung von zwei betonten Einsilbern erzwingt die Kulminativität der prosodischen Prominenz gewissermaßen rhythmische Alternanz durch Hervorhebung einer Silbe durch Wort- oder Phrasenakzent. Gleichzeitig wird die prosodische Parallelität der betonten Einsilber nicht aufgegeben. Denn die Akzentuierung der einen Silbe bedeutet nicht, dass die Betonung der anderen Silbe gelöscht würde - auf der Ebene der lexikalischen Betonung wird das Parallelitätsgebot also eingehalten.

4 Überblick über die Schriften

Die Manuskripte, die Teil dieser kumulativen Habilitationsschrift sind, machen sich diverse empirische und formale Methoden der modernen Linguistik zunutze; es werden sowohl grammatische als auch psycholinguistische Aspekte beleuchtet. Die Arbeiten lassen sich daher auch nicht einem bestimmten Theoriegebäude zuordnen. Obwohl die berichteten Studien nicht direkt aufeinander Bezug nehmen, behandeln sie alle im weitesten Sinne Fragen der phonologischen Identität und ihrer Vermeidung. Die Arbeiten befassen sich dabei schwerpunktmäßig mit prosodischen Domänen oberhalb der Silbe; sie liegen damit in einem Feld, das in Abschnitt 3 bereits umrissen wurde. Die oben in Abschnitt 2 besprochenen Phänomene stellen ergänzend einige Phänomene der phonologischen Identität und ihrer Vermeidung auf der segmentalen Ebene dar.

Bei der Vorstellung der Arbeiten richte ich mich nach dem Kriterium der phonologischen Identität bzw. ihrer Vermeidung. In Abschnitt 4.1 kommen Studien zur Sprache, die sich mit der Präferenz für rhythmische Alternanz von betonten und unbetonten Silben und mit einem möglichen Einfluss des Alternanzgebots auf die Satzverarbeitung befassen.

Studien, die die Wirkung des Parallelismus in der phonologischen Repräsentation diskutieren, werden in Abschnitt 4.2 vorgestellt. In den in Abschnitt 4.3 vorgestellten Studien diskutiere ich die Interaktion von prosodischem Parallelismus und segmentaler bzw. rhythmischer Alternanz. Eine "musikalische Zugabe" (Abschnitt 4.4) schließt die Studienreihe ab.

4.1 Alternanz

4.1.1 Kentner & Vasishth 2016: Prosodic focus marking in silent reading – effects of discourse context and rhythm

Die in Kentner and Vasishth (2016) berichteten Experimente gehen der Frage nach, ob und inwieweit Leser sich bei der syntaktischen Analyse des Lesetextes auf Merkmale der prosodischen Struktur stützen. Wir untersuchen konkret Effekte der *stress clash*-Vermeidung, indem wir in der Umgebung einer syntaktischen Ambiguität, die Anbindung des variabel akzentuierbaren Fokusoperators *auch* betreffend (37), systematisch die lokale Betonungsstruktur variieren. Die zwei Leseexperimente zeigen übereinstimmend, dass Leser Interpretationen vermeiden, die eine bei Assoziation mit Subjekt-Fokus notwendige Akzentuierung des Fokusoperators erfordern (37-b) – und zwar insbesondere dann, wenn die dem ambigen *auch* vorangehende Silbe Betonung trägt. Wir interpretieren dieses Ergebnis als Nachweis der Neigung zu rhythmischer Alternanz von betont und unbetont. Die Experimente belegen damit, dass auch beim stillen Lesen eine prosodische Repräsentation aufgebaut wird, und dass das Prinzip der rhythmischen Alternanz syntaktisch wirksam ist, obwohl die Betonungsstruktur in der schriftlichen Modalität nicht explizit kodiert ist.

(37) a. Objektfokus
*..., dass {Hans, Jochen} auch {*LEHR*linge, Kol*LE*gen} überwacht.*
b. Subjektfokus
..., dass {Hans, Jochen} AUCH *{Lehrlinge, Kollegen} überwacht.*

Dass die Erwartung rhythmischer Alternanz beim Lesen Einfluss auf die syntaktische Verarbeitung des Textes hat, legt nahe, dass Leser nicht nur Rezipienten der im Text gegebenen linguistischen Information sind. Stattdessen scheint es so, dass das Lesen und Verarbeiten von geschriebenen Texten auch ein gewisses Maß an Sprachproduktion beinhaltet, da der Leser dem geschriebenen Text offenbar eine prosodische Struktur zuweisen muss.

Diese Erkenntnis motiviert die Frage, ob auch in der mündlichen Sprachproduktion rhythmisch-prosodische Prinzipien auf die syntaktische Enkodierung einwirken können. Diese Frage greife ich in der folgenden Studie auf.

4.1.2 Kentner & Franz 2018: Rhythmic effects on syntactic encoding are restricted to subsentential projections

Kentner and Franz (2018) untersuchen rhythmische Einflüsse auf die Wahl zwischen eingeleiteten und uneingeleiteten Komplementsätzen. Unter der Annahme, dass der Komplementierer *dass* in der Regel unbetont ist, sollte er im Sinne des Prinzips der rhythmischen Alternanz vermieden werden, wenn die Silben direkt vor und nach der Nebensatzgrenze bereits unbetont sind. In dem Fall sollten Sprecher uneingeleitete Verb-Zweit-Strukturen (38-b) bevorzugen. Umgekehrt sollten mit unbetontem *dass* eingeleitete Komplementsätze (38-a) dann bevorzugt werden, wenn die Nachbarsilben betont sind und daher das Fehlen von *dass* einen *stress clash* nach sich ziehen würde.

(38) a. *Tim {glaubt, glaubte}, dass {Nadja, Nadine} Briefe schreibt.*
b. *Tim {glaubt, glaubte}, {Nadja, Nadine} schreibt Briefe.*

In zwei Sprachproduktionsexperimenten und einer Untersuchung eines spontansprachlichen Korpus' konnten wir keine Belege für diese Hypothese finden. Das rhythmische Alternanzgebot hat also in der mündlichen Sprachproduktion keine so starke Wirkung, dass es die syntaktische Struktur des Komplementsatzes signifikant beeinflussen würde. Wir konnten lediglich in einem schriftsprachlichen Korpus Evidenz in Richtung unserer Hypothese finden: Nach auf schwa-Silbe, also unbetont auslautenden Matrixverben waren signifikant mehr uneingeleitete Komplementsätze zu finden; der unbetonte Komplementierer scheint nach unbetonter Silbe suboptimal. Insgesamt legt dieses Experiment nah, dass es nur einen sehr begrenzten phonologischen Einfluss auf die Satzstrukturierung in der Sprachproduktion gibt; wir haben daher keinen Grund, von klassischen Modellen der Sprachproduktion abzurücken, die davon ausgehen, dass die syntaktische Enkodierung der phonologischen vorangeht und daher nur ein geringer phonologischer Einfluss auf die syntaktische Strukturierung zu erwarten ist (Levelt, 1993).[3] In einer Synopse zeigen wir, dass dort, wo rhythmische Effekte auf die syntaktische Enkodierung nachgewiesen wurden, die betroffenen Strukturen – anders als in unseren Experimenten, die die Struktur der CP betreffen – auf subsententieller Ebene angesiedelt sind; dies trifft übrigens auch auf die syntaktischen Ambiguität zu, die in dem Leseexperiment von Kentner and Vasishth (2016) diskutiert wird. Wir vermuten daher, dass syntaktische Entscheidungen höherer Ordnung von rhythmisch-phonologischen Einflüssen unberührt bleiben.

4.2 Parallelismus

4.2.1 Kentner 2015: Problems of prosodic parallelism: A reply to Wiese & Speyer 2015

In meiner Antwort auf Wiese and Speyer (2015) diskutiere ich drei Probleme, die ich bei Wiese & Speyers Anwendung des Parallelismusgebots auf die Distribution des optionalen schwa in Beispielen wie *Tür* ∼ *Türe* sehe. Ihre aus schriftsprachlichen Korpora entnommenen Daten zeigen, dass die prosodische Struktur des vorangehenden Determinierers (im Fall von *Tür(e)* einsilbiges *die* bzw. trochäisches *eine*) Einfluss auf das Auftreten des optionale Schwa am Nomen hat: prosodisch parallele Strukturen wie *die Tür* und *eine Türe* sollen demnach gegenüber prosodisch asymmetrischen Det-N-Verbindungen bevorzugt werden.

Das erste Problem bei Wiese und Speyers Analyse sehe ich in der Vernachlässigung von Parallelismuseffekten in höheren prosodischen Domänen, die – wie ich zeige – mit Parallelismuseffekten niedrigerer Domänen interagieren können. Zweitens ist die Analyse nur unter der Voraussetzung gültig, dass die Determinierer tatsächlich Füße projizieren – eine Annahme, die aufgrund von bei Funktionswörtern verbreiteten Reduktionsphänomenen fragwürdig ist. Schließlich ist Wiese und Speyers Analyse unvereinbar mit der Beobachtung, dass Determinierer im Phrasen- oder Satzkontext durchaus auch Enklisen zu vorangehenden betonten Silben sein können (39), sodass Determinierer und Nomen (*'ne Flasche*) nicht mehr Teil einer prosodischen Einheit sind, für die das Parallelismusgebot gültig wäre.

(39) *(Hol mir) (mal 'ne) (Flasche) (Bier)*

[3] Siehe auch Golston (1995) and Anttila (2016) für ähnliche Überlegungen über die Interaktion von Syntax und Phonologie in der Grammatik.

Möglicherweise lässt sich die Wirkung des Parallelismusgebots besser anhand der Präferenz für symmetrisch aufgebaute Namen nachweisen – diese sind nämlich weniger anfällig für Reduktionen, und ihre Bestandteile werden in aller Regel nicht durch Phrasengrenzen voneinander getrennt. Ich gehe dieser Vermutung in der folgenden Studie nach.

4.2.2 Kentner 2016: New evidence for prosodic parallelism in German

Anhand eines Korpus' von Bezeichnungen für Musikgenres (häufig handelt es sich um Neuschöpfungen mit englischen Wurzeln) zeige ich in Kentner (2016), dass Folgen von zwei monosyllabischen Füßen oder Wörtern (*hard pop, speedcore*) oder von zwei Trochäen (*finnish indie, psychobilly*) gegenüber asymmetrisch aufgebauten Namen (*swedish pop, death metal*) offenbar deutlich überrepräsentiert sind. Die Neigung zu parallel aufgebauten Strukturen gilt aber anscheinend nur für die Verknüpfung von Fußstrukturen, die im nativen Wortschatz vorkommen, also für die Verknüpfung von Einsilbern und Trochäen. Für komplexere Strukturen mit Dreisilbern habe ich keine Evidenz für prosodischen Parallelismus gefunden.

4.2.3 Kentner & Féry 2013: A new approach to prosodic grouping

Kentner and Féry (2013) untersuchen – dem Beispiel von Wagner (2005) folgend – die Phrasierung von unterschiedlich tief eingebetteten koordinierten Namen wie in (40) und finden dabei Evidenz für prosodischen Parallelismus auf höherer Ebene: Unsere Daten zeigen, dass Sprecher unabhängig von der Verzweigungsrichtung und der Komplexität der Strukturen prosodische Phrasierungen wählen, die die Dauer von Phrasen auf derselben syntaktischen Einbettungstiefe angleicht. Wenn also wie in (40-b) der einfache Name *Nina* auf derselben syntaktischen Einbettungsebene liegt wie die Phrase *Manu und Sonja*, so wird der Name *Nina* prosodisch gelängt, um ihn der Länge der komplexen Schwesterkonstituente anzugleichen – entsprechendes gilt für den Namen *Manu* in (40-c), der syntaktische Schwester zur Phrase *Susi und Nina* ist.

(40) a. *(Susi und Nina) oder (Manu und Sonja)*
b. *Susi oder (Nina oder (Manu und Sonja))*
c. *((Susi und Nina) oder Manu) oder Sonja*

Kentner and Féry (2013) erfassen diese Beobachtung mithilfe eines Similaritätsgebots (41), das im Fall der koordinierten Namen direkt auf der Grundlage der syntaktischen Struktur operiert und die prosodische Anpassung der einzelnen Namen bewirkt.

(41) Die prosodische Grenze nach einer einfachen Konstituente x wird (durch

Längung von x) verstärkt, wenn die Schwesterkonstituente von x komplex ist.

Ein Vergleich von statistischen Modellen zeigt, dass (41) zwar geringeren Einfluss auf die prosodische Wiedergabe der syntaktischen Struktur hat als andere Schnittstellenbedingungen, das Similaritätsgebot gleichwohl einen signifikanten Beitrag zur Erklärung der prosodischen Struktur liefert.

4.2.4 Kentner 2007: Length, ordering preference and intonational phrasing – evidence from pauses

Eine Verletzung des Parallelismusgebots kann zu prosodischer Komplexität führen. In Kentner (2007) berichte ich über ein Experiment, in dem die Phrasierung von koordinierten Nominalphrasen ungleicher Länge in der gelesenen Sprache untersucht wird (42).

(42) *Karl malt ...*
  a. *(den Fluss) und (das winzige geklinkerte Gartenhaus).*
  b. *(das winzige geklinkerte Gartenhaus) und (den Fluss).*

Die Abfolge (42-b) verletzt neben dem Parallelismusgebot auch das sogenannte *Gesetz der wachsenden Glieder* (Behaghel, 1909). Letzteres besagt, dass bei ungleicher Länge das phonologisch leichtere Element bevorzugt dem schweren vorangeht; vgl. Ryan (2017) für einen rezenten Überblick über dieses Gesetz und seine moderneren Reinkarnationen. In den Daten zeigt sich, dass insbesondere die Abfolge *lang vor kurz* (42-b) ein vermehrtes Einsetzen von Phrasengrenzen zwischen den Konjunkten zur Folge hat und damit eine prosodisch deutlich komplexere Struktur.

4.3 Interaktion von Alternanz und Parallelismus

4.3.1 Kentner 2017: On the emergence of reduplication in German morphophonology

Die Reim- und Ablautreduplikationen im Deutschen (43) sind ein Beispiel für das Miteinander von segmental-phonologischer Alternanz und prosodischem Parallelismus.

(43) a. *Mischmasch, Hickhack, Krimskrams*
  b. *Schickimicki, Doppelmoppel, Hasepase*

Als ein besonderer Fall von prosodischer Morphologie richtet sich die Struktur dieser Wörter nach einer optimalen prosodisch-phonologischen Gestalt, die allerdings im Unterschied zu anderen Fällen prosodischer Morphologie – wie die i-Bildung

(Féry, 1997), oder die Beispiele in (32) – nicht auf eine bestimmte Silbenzahl fixiert ist, sondern sowohl Zweisilber als auch Viersilber erlaubt. Für die reduplikativen Zwei- und Viersilber gilt, dass sie prosodisch parallel im Sinne von (36) aufgebaut sind – eine ungerade Silbenzahl verbietet sich deshalb. In meiner Studie grenze ich Reim- und Ablautreduplikation von anderen, auf den ersten Blick ähnlichen Strukturen ab, und schlage eine optimalitätstheoretische Grammatik vor, die im Stande ist, beide Formen der Reduplikation zu generieren. Für Reim- und Ablautreduplikation gilt, dass Basis und Reduplikant segmental unterschieden sind. Der Unterschied muss ist allerdings minimal sein, insofern er in der Regel nicht mehr als ein Segment betragen darf. Vermutlich ergibt sich diese Tatsache aus der Eigenschaft der Reduplikation, Material aus der Basis zu kopieren und der oben beschriebenen, generellen Beschränkung, wonach adjazente phonologische Einheiten segmental unterschiedlich sein sollen (s. Abschnitt 1 und insbesondere 1.7).

Anhand eines größeren Korpus' und einer Fragebogenstudie weise ich nach, dass Reim- und Ablautreduplikation im Deutschen unter den genannten Bedingungen produktive Wortbildungsmuster sind.

4.3.2 Kentner 2018: Schwa-optionality and the prosodic shape of words and phrases

In dieser Studie stelle ich drei Experimente vor, anhand derer ich den Einfluss des Prinzips der rhythmischer Alternation einerseits und des Parallelismusgebots andererseits auf morphophonologische und morphosyntaktische Variation im Deutschen diskutiere. Konkret untersuche ich das Auftreten der optionalen Schwa-Silbe in Adverbien wie *gern(e), selbst/selber, meist/meistens, lang(e)* unter verschiedenen prosodischen Bedingungen, die durch unterschiedliche flektierte Formen des auf das Adverb folgenden Verbs (*tun, getan* oder *machen, gemacht*) repräsentiert werden. In einem schriftlichen Korpus zeigt sich, dass rhythmisch alternierende Adverb-Verb-Abfolgen gegenüber unrhythmischen Abfolgen mit *stress clash* (*gern tun*) oder *stress lapse* (*gerne gemacht*) bevorzugt werden. In dieser Umgebung gab es keinen signifikanten Hinweis auf die Wirksamkeit des prosodischen Parallelismus. Ich schließe daraus, dass das Prinzip der rhythmischen Alternanz einen stärkeren Einfluss auf morpho-phonologische Gegebenheiten hat als das Parallelismusgebot. In einem Leseexperiment zeigt sich allerdings ein – wenn auch schwacher – Parallelismuseffekt: Probanden setzen das optionale Schwa bei Wörtern wie *gern(e)* tendentiell so ein, dass das Adverb dieselbe prosodische Struktur erhält wie ein vorangehendes Nomen; nach trochäischem Zweisilber wird das Schwa am Adverb deutlich häufiger eingesetzt (*...will im Garten gerne...*), als nach betontem und akzentuiertem Einsilber; in letzterem Fall wird monosyllabisches *gern* bevorzugt (*...will im Hof gern...*), d.h. viele Probanden lassen anscheinend einen *stress clash* zugunsten einer prosodisch parallelen Phrasierung zu. Allerdings stelle ich zur Diskussion, ob das Adverb in dieser Konstellation nicht möglicherweise völlig deakzentuiert wird und in einen Fuß mit dem monosyllabischen *Hof* integriert wird. Für diesen Fall wäre wiederum rhythmische Alternanz als prosodisches Strukturprinzip ausschlaggebend. Die Frage der prosodischen Repräsentation bleibt an dieser Stelle offen.

Die Neigung zu rhythmischer Alternanz zeigt sich auch in der Wahl zwischen postnominalem Genitiv (*der Knopf der Hose*) und der quasi-synonymen Präpositionalkonstruktion mit *von* (*der Knopf von der Hose*). Die Präpositionalkonstruktion hat grundsätzlich eine unbetonte Silbe mehr zwischen Bezugsnomen und Attribut als die Struktur mit Genitiv; sie läuft damit dem Prinzip der rhythmischen Alternanz eher zuwider, als die Konstruktion mit Genitiv. Tatsächlich zeigt sich, dass Probanden in einer *forced-choice*-Aufgabe die Präpositionalkonstruktionen bei Bezugsnomen mit unbetonter Ultima (und damit mindestens drei unbetonten Silben in Folge: *...Knöpfe von der...*) zugunsten des Genitivs vermeiden – ein Umstand den ich im Sinne des Prinzips der rhythmischen Alternanz interpretiere.

4.4 Musikalische Zugabe

In den vorangehenden Abschnitten war von lautlicher Identität und Identitätsvermeidung in unterschiedlichen segmentalen und prosodischen Domänen die Rede. Die im Folgenden dargestellte Studie macht sich eine andere Perspektive auf lautliche Identität zu eigen, indem sie die sprachspezifische Verarbeitung phonetisch identischer Stimuli in den Mittelpunkt der Untersuchung stellt. Der Fokus der Untersuchung liegt auf sprachrhythmischen Prominenzen und ihrer Rolle in der Segmentierung des Stimulusmaterials.

4.4.1 Kentner 2015b: Rhythmic segmentation in auditory illusions: Evidence from cross-linguistic mondegreens

Hörer sind regelmäßig mit Liedern konfrontiert, deren Text in einer Sprache verfasst ist, die nicht ihre Muttersprache ist. Dabei kommt es vor, dass Hörer Fragmente in den Liedtexten als ihrer eigenen Muttersprache zugehörig identifizieren, obwohl sie sich der Tatsache bewußt sind, dass das Lied in einer anderen Sprache vorgetragen wird (44).

(44) *Hope of deliverance* engl. Originaltext [Paul McCartney]
Hau auf die Leberwurst deutsches Perzept

In dieser Studie untersuche ich solche sprachübergreifenden Verhörer. Ich beschränke mich dabei auf ein eigens zusammengestelltes Korpus von deutsch bzw. französisch wahrgenommenen Liedverhörern, deren Originaltext jeweils Englisch ist. Mithilfe dieses Materials vergleiche ich die Wortsegmentierung des Originaltexts mit dem Verhörertext. Der Vergleich zeigt, dass sowohl für deutsche als auch französische Verhörer rhythmische Prominenzen für die Wortsegmentierung ausschlaggebend sind. Allerdings wird deutlich, dass rhythmische Prominenzen bei Verhörern von Sprechern des Deutschen vor allem auf einer wortinitialen Silbe liegen, während französische Verhörer wort- oder phrasenfinale Prominenzen aufweisen. Dieser Unterschied in der Rolle von rhythmischen Prominenzen für die Wortsegmentierung entspricht der Rolle der Wort- und Phrasenprosodie in diesen Sprachen. Die Betonung im Deutschen ist in der Regel fußinitial und damit auch – weil die meisten Wörter des Kernwortschatzes Einsilber oder Trochäen sind – wortinitial. Das Französische dagegen hat keine distinktive Wortbetonung, prosodisch prominente Silben tragen Grenztöne und markieren damit Phrasenfinalität. Die Verhörer belegen somit, dass je nachdem, welche Rolle die rhythmische Prominenz in der Einzelsprache hat, lautlich identische Texte unterschiedliche phonologische Repräsentationen hervorrufen.

Literatur

Anttila, A., M. Adams, and M. Speriosu. 2010. The role of prosody in the English dative alternation. *Language and Cognitive Processes* 25:946–981.

Anttila, Arto. 2016. Phonological effects on syntactic variation. *Annual Review of Linguistics* 2:115–137.

Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25:110.

Breen, Mara, Duane G Watson, and Edward Gibson. 2011. Intonational phrasing is constrained by meaning, not balance. *Language and Cognitive Processes* 26:1532–1562.

Brugmann, Karl. 1917. Gleichklangvermeidung in der lautgesetzlichen Entwicklung und in der Wortbildung. *Indogermanische Forschungen* 38:117–128.

Cooper, William E, and Jeanne Paccia-Cooper. 1980. *Syntax and speech*. Harvard University Press.

Domahs, Ulrike, Wolfgang Kehrein, Johannes Knaus, Richard Wiese, and Matthias Schlesewsky. 2009. Event-related potentials reflecting the processing of phonological constraint violations. *Language and Speech* 52:415–435.

Dressler, Wolfgang U, Katarzyna Dziubalska-Kołaczyk, Natalia Gagarina, and Marianne Kilani-Schoch. 2005. Reduplication in child language. In *Studies on reduplication*, ed. Bernhard Hurch, 455–474. Walter de Gruyter.

Eisenberg, Peter. 1991. Syllabische Struktur und Wortakzent. Prinzipien der Prosodik deutscher Wörter. *Zeitschrift für Sprachwissenschaft* 10:37–64.

Ferreira, Fernanda. 2002. Prosody. In *Encyclopedia of Cognitive Science*, ed. Lynn Nadel. London: Macmillan Reference Ltd.

Féry, Caroline. 1997. Uni und Studis: die besten Wörter des Deutschen. *Linguistische Berichte* 172:461–489.

Fikkert, Paula, and Claartje Levelt. 2008. How does Place fall into Place? In *Contrast in phonology: Theory, perception, acquisition*, ed. Peter Avery, B Elan Dresher, and Keren Rice, 231–268. Berlin: Walter de Gruyter.

Finkbeiner, Rita. 2012. Naja, normal und normal. Zur Syntax, Semantik und Pragmatik der x-und-x-Konstruktion im Deutschen. *Zeitschrift für Sprachwissenschaft* 31:1–42.

Finkbeiner, Rita. 2014. Identical constituent compounds in German. *Word Structure* 7:182–213.

Fodor, J.D. 1998. Learning to parse? *Journal of Psycholinguistic Research* 27:285–319.

Frazier, Lyn, and Janet Dean Fodor. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6:291–325.

Freywald, Ulrike. 2015. Total reduplication as a productive process in German. *Studies in Language* 39:905–945.

Frisch, Stefan A, Janet B Pierrehumbert, and Michael B Broe. 2004. Similarity avoidance and the OCP. *Natural Language & Linguistic Theory* 22:179–228.

Gee, James Paul, and François Grosjean. 1983. Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive psychology* 15:411–458.

Ghini, M. 1993. Φ-formation in Italian: A new proposal. *Toronto Working Papers in Linguistics* 41–78.

Golston, Chris. 1995. Syntax outranks phonology: evidence from Ancient Greek. *Phonology* 12:343–368.

Gordon, P. C., Barbara J. Grosz, and Laura A. Gilliom. 1993. Pronouns, names and the centering of attention in discourse. *Cognitive Science* 17(3):311–348.

Grosjean, Francois, Lysiane Grosjean, and Harlan Lane. 1979. The patterns of silence: Performance structures in sentence production. *Cognitive psychology* 11:58–81.

Hardcastle, William J, and Nigel Hewlett. 2006. *Coarticulation: Theory, data and techniques*. Cambridge: Cambridge University Press.

Hayes, B. 1995. *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Henrich, Karen. 2015. It´s all about the rhythm: A neurocognitive approach towards the Rhythm Rule in German and English. Doctoral Dissertation, Philipps-Universität Marburg, Marburg.

Hyman, Larry M. 2006. Word-prosodic typology. *Phonology* 23:225–257.

Kentner, G. 2007. Length, ordering preference and intonational phrasing: Evidence from pauses. In *Proceedings of Interspeech 2007*. Antwerp.

Kentner, Gerrit. 2015. Problems of prosodic parallelism: A reply to Wiese & Speyer (2015). *Linguistics* 53:1233–1241.

Kentner, Gerrit. 2016. New evidence for prosodic parallelism in German. In *Tagungsband Phonetik und Phonologie 12*, ed. Christoph Draxler and Felicitas Kleber, 88–89. München.

Kentner, Gerrit. 2017. On the emergence of reduplication in German morphophonology. *Zeitschrift für Sprachwissenschaft* 36:233–277.

Kentner, Gerrit. 2018. Schwa-optionality and the prosodic shape of words and phrases. In *Empirical approaches to the phonological structure of words*, ed. Christiane Ulbrich, Alexander Werth, and Richard Wiese, xxx–xxx. Berlin: de Gruyter.

Kentner, Gerrit, and Caroline Féry. 2013. A new approach to prosodic grouping. *The Linguistic Review* 30:277–311.

Kentner, Gerrit, and Isabelle Franz. 2018. Rhythmic effects on syntactic encoding are restricted to subsentential projections. *Ms. Goethe-Universität Frankfurt* xx:xx.

Kentner, Gerrit, and Shravan Vasishth. 2016. Prosodic focus marking in silent reading: effects of discourse context and rhythm. *Frontiers in Psychology* 7:319.

Kiparsky, Paul. 1966. Über den deutschen Akzent. *Studia Grammatica* VII:69–98.

Kloeke, WUS van Lessen. 1982. *Deutsche phonologie und morphologie: Merkmale und markiertheit*, volume 117. Walter de Gruyter.

Köbler, Gerhard. 2007. *Historisches lexikon der deutschen länder: die deutschen territorien vom mittelalter bis zur gegenwart*. Frankfurt: CH Beck.

Leben, William Ronald. 1973. Suprasegmental phonology. Doctoral Dissertation, Massachusetts Institute of Technology.

Levelt, Willem JM. 1993. *Speaking: From intention to articulation*. MIT press.

Macfarland, Talke, and Janet Pierrehumbert. 1991. On ich-laut, ach-laut and structure preservation. *Phonology* 8:171–180.

Mayer, Thomas, Christian Rohrdantz, Frans Plank, Peter Bak, Miriam Butt, and Daniel A Keim. 2010. Consonant co-occurrence in stems across languages: Automatic analysis and visualization of a phonotactic constraint. In *Proceedings of the 2010 Workshop on NLP and Linguistics: Finding the Common Ground*, 70–78. Association for Computational Linguistics.

McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12:373–418.

McCarthy, John J. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17:207–263.

Merchant, Jason. 1996. Alignment and fricative assimilation in german. *Linguistic Inquiry* 27:709–719.

Müller, Gereon. 1997. Beschränkungen für Binomialbildung im Deutschen: Ein Beitrag zur Interaktion von Phraseologie und Grammatik. *Zeitschrift für Sprachwissenschaft* 16:5–51.

Müller, Gereon. 2016. Predicate doubling by phonological copying. In *Linguistische arbeitsberichte 93*, ed. Katja Barnickel, Matías Guzmán Naranjo, Johannes Hein, Sampson Korsah, Andrew Murphy, Ludger Paschen, Zorica Puškar, and Joanna Zaleska, 127–148. Leipzig: Universität Leipzig.

Myrberg, Sara. 2013. Sisterhood in prosodic branching. *Phonology* 30:73–124.

Noske, Manuela. 1997. Feature spreading as dealignment: the distribution of [ç] and [x] in German. *Phonology* 14:221–234.

Ortmann, Wolf Dieter. 1993. *Kernmorpheme im deutschen: 4561 allgemeinsprachliche deutsche morphemgestalten im vergleich zwischen zehn quellen*. Goethe-Institut.

Pareyon, Gabriel. 2011. *On musical self-similarity: Intersemiosis as synecdoche and analogy*. Semiotic Society of Finland.

Pater, Joe, and Adam Werle. 2003. Direction of assimilation in child consonant harmony. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 48:385–408.

Pierrehumbert, Janet. 1993. Dissimilarity in the arabic verbal roots. In *Proceedings of the North East Linguistics Society*, volume 23, 367–381.

Plag, Ingo. 1998. Morphological haplology in a constraint-based morphophonology. In *Phonology and morphology of the Germanic languages*, ed. Richard Wiese and Wolfgang Kehrein, 199–215. Tübingen: Niemeyer.

Pozdniakov, Konstantin, and Guillaume Segerer. 2007. Similar place avoidance: A statistical universal. *Linguistic Typology* 11:307–348.

Riad, Tomas. 2012. Culminativity, stress and tone accent in central swedish. *Lingua* 122:1352–1379.

Ryan, Kevin. 2017. *Prosodic end-weight reflects phrasal stress*. Harvard University: Manuscript.

Saba Kirchner, Jesse. 2010. Minimal reduplication. Doctoral Dissertation, University of California Santa Cruz, CA.

Schein, Barry, and Donca Steriade. 1986. On geminates. *Linguistic Inquiry* 17:691–744.

Schlüter, J. 2005. *Rhythmic grammar: the influence of rhythm on grammatical variation and change in English*. Berlin: Mouton de Gruyter.

Schweitzer, Katrin, Kerstin Eckart, Petra Augurzky, Natalie Lewandowski, and Grzegorz Dogil. 2011. A corpus-phonetic investigation of prosodic balance in German. In *Proceedings of AMLaP 2011*. Paris.

Selkirk, E. 2000. The interactions of constraints on prosodic phrasing. In *Prosody: Theory and experiment*, ed. M. Horne, 231–261. Dordrecht: Kluwer Academic.

Selkirk, E.O. 1984. *Phonology and syntax*. MIT Press Cambridge, Mass.

Shih, Stephanie, Jason Grafmiller, Richard Futrell, and Joan Bresnan. 2015. Rhythm's role in genitive construction choice in spoken English. In *Rhythm in Cognition and Grammar. A Germanic Perspective*, ed. Ralf Vogel and Ruben van de Vijver, 207–233. Berlin: de Gruyter.

Vogel, Ralf, Ruben van de Vijver, Sonja Kotz, Anna Kutscher, and Petra Wagner. 2015. Function words in rhythmic optimisation. In *Rhythm in Cognition and Grammar. A Germanic Perspective*, ed. Ralf Vogel and Ruben van de Vijver, 255–276. Berlin: de Gruyter.

Vogt, Barbara. 2013. Consequences of rhythmic secondary stress in German: stress patterns in language games and past participle formation. *Linguistische Berichte* 2013:171–191.

Wagner, Michael. 2005. Prosody and recursion. Doctoral Dissertation, Massachusetts Institute of Technology.

Wegener, Heide. 2004. Pizzas und Pizzen–die Pluralformen (un)assimilierter Fremdwörter im Deutschen. *Zeitschrift für Sprachwissenschaft* 23:47–112.

Wiese, Richard. 1996. *The phonology of German*. Clarendon Press Oxford.

Wiese, Richard. 2001. The structure of the german vocabulary: edge marking of categories and functional considerations. *Linguistics* 39:95–116.

Wiese, Richard. 2016. Prosodic parallelism - comparing spoken and written language. *Frontiers in Psychology* 7:1598.

Wiese, Richard, and Augustin Speyer. 2015. Prosodic parallelism explaining morphophonological variation in German. *Linguistics* 53:525–559.

Wurzel, Wolfgang Ullrich. 1970. *Studien zur deutschen Lautstruktur*. Berlin: Akademie-Verlag.

Yip, Moira. 1988. The obligatory contour principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19:65–100.

5 Anhang

Dieser Anhang enthält die im Rahmen dieser Arbeit eingereichten Schriften. Drei davon sind gemeinsam mit Ko-Autoren entstanden. Für diese Arbeiten weise ich im Folgenden den eigenen Beitrag aus.

5.1 Eigener Beitrag bei gemeinschaftlich verfassten Arbeiten

Die Arbeit mit dem Titel "A new approach to prosodic grouping" (Kentner and Féry, 2013) habe ich gemeinsam mit Caroline Féry verfasst. Die Idee für das dort berichtete Sprachproduktionsexperiment stammt von ihr und es wurde unter ihrer Ägide durchgeführt. Ich war zuständig für die akustisch-phonetische Analyse der Produktionsdaten und die inferenzstatistische Auswertung. Auf Grundlage des Produktionsexperiments habe ich das Perzeptionsexperiment konzipiert, durchgeführt und ausgewertet. Die Interpretation der Daten und das Verfassen des Textes war eine Gemeinschaftsleistung, wobei mein Beitrag insbesondere in der Formulierung des *Proximity/Similarity*-Modells und dem statistischen Vergleich mit anderen Modellen der Syntax-Phonologie-Schnittstelle lag.

Mit Shravan Vasishth habe ich die Arbeit "Prosodic focus marking in silent reading – effects of discourse context and rhythm" veröffentlicht. Ich habe die dort berichteten Experimente konzipiert, durchgeführt und ausgewertet, die Daten interpretiert und den Großteil des Manuskripts erstellt. Shravan Vasishth hat Laborinfrastruktur zur Verfügung gestellt und war für die im Anhang der Arbeit dargestellte statistische Analyse nach Bayes zuständig.

Das Manuskript mit dem Titel "Rhythmic effects on syntactic encoding are restricted to subsentential projections" ist zusammen mit Isabelle Franz entstanden. Ich habe die dort berichteten Experimente konzipiert, die Daten statistisch ausgewertet und interpretiert sowie den Großteil des Manuskripts verfasst. Isabelle Franz hat das Material für die Sprachproduktionsexperimente erstellt, die Experimente durchgeführt, interpretiert und die entsprechenden Methodenteile des Manuskripts verfasst.

frontiers in Psychology

ORIGINAL RESEARCH
published: 08 March 2016
doi: 10.3389/fpsyg.2016.00319

Prosodic Focus Marking in Silent Reading: Effects of Discourse Context and Rhythm

Gerrit Kentner[1]* *and Shravan Vasishth*[2]

[1] *Department of Linguistics, Goethe University, Frankfurt, Germany,* [2] *Department of Linguistics, University of Potsdam, Potsdam, Germany*

Understanding a sentence and integrating it into the discourse depends upon the identification of its focus, which, in spoken German, is marked by accentuation. In the case of written language, which lacks explicit cues to accent, readers have to draw on other kinds of information to determine the focus. We study the joint or interactive effects of two kinds of information that have no direct representation in print but have each been shown to be influential in the reader's text comprehension: (i) the (low-level) rhythmic-prosodic structure that is based on the distribution of lexically stressed syllables, and (ii) the (high-level) discourse context that is grounded in the memory of previous linguistic content. Systematically manipulating these factors, we examine the way readers resolve a syntactic ambiguity involving the scopally ambiguous focus operator *auch* (engl. "too") in both oral (Experiment 1) and silent reading (Experiment 2). The results of both experiments attest that discourse context and local linguistic rhythm conspire to guide the syntactic and, concomitantly, the focus-structural analysis of ambiguous sentences. We argue that reading comprehension requires the (implicit) assignment of accents according to the focus structure and that, by establishing a prominence profile, the implicit prosodic rhythm directly affects accent assignment.

Keywords: linguistic rhythm, focus, accent, reading, implicit prosody, syntactic parsing, sentence comprehension, eye tracking

Edited by:
Matthew W. Crocker, Saarland University, Germany

Reviewed by:
Christoph Scheepers, University of Glasgow, UK
Richard Wiese, Philipps-Universität Marburg, Germany

****Correspondence:***
Gerrit Kentner
kentner@lingua.uni-frankfurt.de

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *12 October 2015*
Accepted: *19 February 2016*
Published: *08 March 2016*

Citation:
Kentner G and Vasishth S (2016) Prosodic Focus Marking in Silent Reading: Effects of Discourse Context and Rhythm. Front. Psychol. 7:319. doi: 10.3389/fpsyg.2016.00319

1. INTRODUCTION

What are the factors determining the syntactic analysis of written text and how do they interact? The vast literature on written sentence comprehension suggests that readers make use of a multitude of information sources in order to extract structure from the printed letter string and compute its meaning. Some of these sources are represented directly in print, e.g., the words that contribute their meanings, or the punctuation that marks the partitioning of phrasal chunks. Other kinds of information have to be derived or inferred from the reader's linguistic and world knowledge. In making such inferences, the reader forms interpretations that constitute predictions about the upcoming text. These predictions may or may not turn out to be compatible with the actual structure of the sentence. The ease with which a reader traverses a text is based to a great extend on how accurate his predictions are.

In this study, we will be concerned with the interaction of two information sources that (i) tap the reader's linguistic knowledge, (ii) have no direct representation in print, and (iii) have each been shown to be influential in the reader's text comprehension process. One of these is the discourse representation (hereafter, context) which is based on the memory of previous linguistic content. The other more local type of information concerns the prosodic structure, specifically the linguistic rhythm that emerges from the succession of lexically strong and weak syllables. The results of two reading experiments presented here attest that discourse context and local linguistic rhythm, two otherwise independent phenomena, conspire to guide the syntactic analysis of structurally ambiguous sentences.

1.1. Implicit Prosody and Discourse Context in Reading

There is hardly any doubt that readers generate a mental prosodic-phonological representation of written texts even in silent reading (Chafe, 1988; Frost, 1998; Ashby and Clifton, 2005; Ashby and Martin, 2008; Savill et al., 2011) and a growing body of evidence supports the idea that these representations, conventionally called *implicit prosody* (Fodor, 2002), co-determine the way in which syntactic ambiguities are resolved (e.g., Bader, 1998; Hirose, 2003; Jun, 2003; Hwang and Steinhauer, 2011), see Breen (2014) for a review. In our own work, we found that readers, when faced with syntactically ambiguous structures, avoid interpretations the phonological representation of which involves a stress clash (i.e., a sequence of two adjacent syllables bearing lexical or post-lexical stress), and instead favor syntactic alternatives that allow for more felicitous, alternating rhythm of strong and weak syllables (Kentner, 2012, 2015; McCurdy et al., 2013). Similarly, Breen and Clifton (2011, 2013) provide evidence for very early prosodic effects contributing significantly to reading effort in ambiguous sentences when the reanalysis of the part-of-speech in noun-verb homographs involves a change in lexical stress [e.g., ABstract (noun) vs. abSTRACT (verb)]. These studies suggest that representations of lexical stress and the expectation of rhythmically alternating syllabic structure not only reflect but potentially direct readers' syntactic parsing decisions.

As for the role of contextual information in sentence comprehension, it has been shown extensively that the previous discourse may bring about strong expectations that guide the syntactic analysis: relevant information in the context may lead to the cancellation of otherwise strong garden path effects (e.g., Altmann and Steedman, 1988; Spivey and Tanenhaus, 1998; Binder et al., 2001; Snedeker and Trueswell, 2004). This has been taken as key evidence in favor of models embodying a multitude of potentially competing information sources simultaneously constraining the way in which the sentence is analyzed (cf. MacDonald et al., 1994; McRae et al., 1998; van Gompel et al., 2001).

Yet, while the influence of both implicit prosody and context on syntactic parsing are each attested, it remains largely unclear whether and how exactly these two kinds of constraint interact in guiding the parsing process. We are aware of two studies that explore the effects of both implicit prosody and discourse context in reading. The first one by Stolterfoht et al. (2007) uses ERP to study the processing of a certain type of ellipsis, so-called replacives (Drubig, 1994), in which a stranded argument is contrastively related to an argument in the preceding main clause (the correlate) (1).

(1) Am Dienstag hat ...
On Tuesday has ...

a. der Rektor (nur) [den SCHÜler]$_F$ getadelt,
the$_{nom}$ principal (only) the$_{acc}$ pupil critizised,
nicht [den LEHrer]$_F$.
not the$_{acc}$ teacher.
On Tuesday, the principal criticized the pupil, not the teacher.

b. (nur) [der REKtor]$_F$ den Schüler getadelt,
(only) the$_{nom}$ principal the$_{acc}$ pupil critizised,
nicht [der LEHrer]$_F$.
not the$_{nom}$ teacher.
On Tuesday, the principal, not the teacher, criticized the pupil.

In the study by Stolterfoht et al. (2007), the morphological case of this stranded, sentence-final argument determines which argument in the main clause is contrastively focussed and, correspondingly, accented—viz. the one bearing the same morphological case. With the presence or absence of the focus particle *nur* ("only"), Stolterfoht et al. (2007) varied the need for the reader to revise the default reading with wide focus to a narrow focus reading, and—depending on the association of *nur* with the subject or the object, they manipulated the need for revising the implicit accent placement when encountering the replacive argument. The ERP results suggest that these two processes—restructuring of focus domains and reassignment of (implicit) accentuation—are independent as they engender different ERP signatures. However, the study remains inconclusive as to whether and how focus structure and implicit accent placement interact in determining a syntactic analysis of the written sentence.

McCurdy et al. (2013) studied the effects of implicit prosody and contextual bias on syntactic parsing using eye-tracking methodology. Building on Bader (1996) and Kentner (2012), the target sentence involves an ambiguity concerning the word sequence *nicht mehr* which could either be resolved as a temporal adverb or as a negated comparative quantifier. In the latter case, *mehr* is accented in a spoken rendition of the sentence, while the temporal reading engenders main phrasal accent on the following verb. In their study, McCurdy et al. (2013) presented readers with a context sentence that was devised so as to bias for one specific reading of the subsequently presented ambiguous target sentence—a manipulation akin to syntactic priming (cf. the boxed portions in the context in (2), corresponding to comparative or temporal adverbs, respectively). As in Kentner (2012), the target sentence was subject to prosodic manipulations concerning the distribution of lexically stressed and unstressed syllables in the ambiguous region; this manipulation lead to either a rhythmically alternating sequences of lexical stresses or to stress clash, depending on the syntactic analysis, which in turn determined the (implicit) accentuation of the ambiguous word *mehr*.

(2) *Context*
COMPARATIVE:
Der Manager verlangt von Peer, länger zu trainieren, als alle anderen.
The manager expects Peer to train longer than all the others.
TEMPORAL:
Peers Manager hat leider schon oft zuviel von Peer gefordert.
Peer's manager has often asked too much of Peer.
Target
Peer denkt, dass der Trainer...
Peer thinks that the trainer...
TEMP... nicht mehr {zulassen/erlauben} sollte, dass er so viel trainiert.
...should not {permit/allow} anymore that he trains so much.
COMP... nicht MEHR {zulassen/erlauben} sollte, als tägliches Training.
...should not {permit / allow} more than daily training.

Replicating findings by Kentner (2012), the results reveal that the readers' avoidance of stress clash configurations significantly contributed to their parsing decisions. This effect was detectable already before the readers' eyes made contact with the disambiguating region. Effects of context on syntactic ambiguity resolution affected later parsing stages only and there was hardly any interaction between these information sources in the eye-movement record.

To summarize, the current state of affairs suggests that, if at all, local linguistic rhythm and more global discourse context interact only weakly, with local prosodic effects preceding any effects of the contextual manipulation. Although contextual information has been reported to affect the earliest stages of sentence comprehension, preferences from more local information sources have been claimed to potentially override contextual biases (Pickering and van Gompel, 2006). This may in fact explain the relatively late influence of context reported in McCurdy et al. (2013). An alternative explanation for the late effect of the contextual manipulation is the non-compelling nature of the bias that was introduced in McCurdy et al. (2013). In contrast to other studies probing contextual influences on syntactic parsing (e.g., Altmann and Steedman, 1988), McCurdy et al. (2013) did not aim at directly manipulating discourse representations. Rather, the context merely anticipated one of the morpho-syntactic structures of the ambiguous target sentence to create a bias for the corresponding interpretation.

1.2. The Prosody and Syntax of the Focus Particle *auch*

To specifically address the interplay of discourse representations and implicit prosody in sentence comprehension, we set out to study a different kind of syntactic ambiguity, the proper resolution of which hinges on contextual information. The ambiguity concerns the interpretation of the focus particle *auch* (engl.: "also") in German (cf. Altmann, 1976; Jacobs, 1983; Sudhoff, 2008; Féry, 2009). Consider the ambiguous example in (3) with the three presuppositional interpretations in (3-a), (3-b), and (3-c). In writing, (3) is ambiguous with respect to the scope of *auch*, which may associate with either subject focus or object or VP-focus[1]. In the oral rendition, the ambiguity is (partly) resolved by prosody: unaccented *auch* and nuclear accent (the most prominent pitch accent in a sentence) on the object *Keller* presupposes that other objects beside the one stated are being rummaged through—hence, *auch* associates with focus on the object *Keller* or on the whole VP *Keller durchstöbert*. The interpretations with object focus and VP-focus have comparable prosodic renderings. Conversely, a rendition with accent on *auch* and a deaccented VP (*Keller durchstöbert*) presupposes that the object, in fact the whole VP, is outside the focus induced by *auch*; this accented rendition of *auch* leads to the presupposition that another person in addition to *Herbert* is the agent of the event expressed in the predicate (*auch* associates with the subject and, consequently, the subject is focussed).

(3) Sonja meint, dass Herbert auch Keller durchstöbert.
lit.: Sonja thinks that Herbert also cellars rummages.
Sonja thinks that Herbert(, too,) is rummaging through cellars(, too).
a. *H. is rummaging through something in addition to cellars.*
(object focus)
b. *H. is doing something in addition to rummaging through cellars.*
(VP focus)
c. *Somebody apart from H. is rummaging through cellars.*
(subject focus)

A preceding context that renders the object in the target sentence either discourse-new or given restricts the room for interpretation to one of the possible interpretations. That is, a context like (4-a) makes the object *Keller* in (3) a new discourse entity. Hence, it is only compatible with *auch* associating with object focus. In that case, nuclear accent is required on the focused object, leaving *auch* prosodically unaccented. In contrast, in (4-b), the whole VP including the object *Keller* is explicitly mentioned. According to the mapping rules concerning information structure and prosody (e.g., Gussenhoven, 1983; Ladd, 1996; Vallduví and Engdahl, 1996; Schwarzschild, 1999; Féry and Samek-Lodovici, 2006; Krifka, 2006; Truckenbrodt, 2006), the givenness of the VP induces its deaccentuation and *auch* becomes the locus of nuclear accent (Féry, 2009),

[1]A fourth reading is available in which *auch* associates with verb focus, skipping a (given) object. The corresponding interpretation presupposes that H. is doing something to cellars in addition to rummaging through them. This reading is highly marked in that *auch* is not adjacent to the focus it associates with. Furthermore, it requires a prosodic rendition that deviates from normal phrasal stress in transitive VPs, i.e., with an unaccented object and nuclear accent on the verb. This reading is not viable in either of the contextual manipulations we devise in the experiment below and it will therefore be disregarded.

thus signaling association with focus on the subject. That is, when preceded by a relevant context, the ambiguity in (3) is properly resolved on the object. Its information status (new or given) unequivocally determines the syntactic association of the focus particle, and, correspondingly, the position of the accent.

(4) Herbert und Karlo sammeln alte Möbel für den Flohmarkt.
Herbert and Karlo collect old furniture for the fleamarket.

- a. Karlo durchstöbert Garagen.
  Karlo rummages through garages.
  Karlo is rummaging through garages.
- b. Karlo durchstöbert Keller.
  Karlo rummages through cellars.
  Karlo is rummaging through cellars.

2. EXPERIMENTS

2.1. General Design

To probe the interaction of implicit prosodic rhythm and contextual information, we applied a a 2 × 2 × 2 factorial design with two rhythmic factors crossed with the above contextual variation, which induces either subject or object focus in the target sentence. First, for the rhythmic context to the left of the ambiguous *auch* (RhythmLeft), the lexical material of the target sentences was constructed to yield a trochaic beat with every other syllable bearing lexical stress. The syllabic structure of the proper name directly preceding *auch* was systematically varied, with either a monosyllable or a disyllabic trochee [contrast between conditions a,b,c,d (trochaic name) vs. e,f,g,h (monosyllabic name) in (5)]. The logic of RhythmLeft is based on evidence for rhythmic entrainment (e.g., Dilley and McAuley, 2008; Niebuhr, 2009; Schmidt-Kassow and Kotz, 2009, w.r.t. auditory linguistic rhythm). If the proper name preceding ambiguous *auch* is trochaic, i.e., ends in an unstressed syllable (conditions a,b,c,d), *auch* falls onto a strong position of the beat established by the preceding word string and is thus more likely to receive prosodic prominence in the form of a (nuclear) accent. Conversely, if preceded by a monosyllabic word, *auch* would be in off-beat position which is predicted to hamper assignment of prosodic prominence.

The rhythmic environment to the right (RhythmRight) is manipulated on the object noun with lexical stress falling either on the initial or onto the second syllable (contrast between conditions a,c,e,g vs. b,d,f,h). An object bearing initial stress leads to a stress clash when the preceding *auch* is prosodically prominent. An iambic object, featuring an unstressed initial syllable, leads to a stress lapse when *auch* remains unaccented.

That is, depending on the accentuation of *auch* as determined by the discourse context, the rhythmic manipulations lead to alternating sequences of stressed and unstressed syllables or to phonologically unsatisfactory clashes or lapses in the context of *auch*. On the basis of our previous studies (Kentner, 2012; McCurdy et al., 2013), we assume that readers favor syntactic parses whose phonological representation has a favorable rhythm. Reading difficulties are predicted to emerge when the contextual manipulation forces a syntactic parse with rhythmically deviant prosodic structure.

(5) I Sonja Kohn und Herbert Otten sind bei einer Sicherheitsfirma angestellt.
Klaus hat erfahren, dass Sonja Kohn Kollegen überwacht.
Sonja Kohn and Herbert Otten work for a security company.
Klaus has learned that Sonja Kohn supervises colleagues.

- a Carla glaubt, dass Herbert Otten auch Kollegen überwacht.
  Carla thinks that Herbert Otten supervises colleagues, too.
  (SubjFoc; RhythmL=on beat; RhythmR=no Clash)
- b Carla glaubt, dass Herbert Otten auch Lehrlinge überwacht.
  Carla thinks that Herbert Otten supervises apprentices, too.
  (ObjFoc; RhythmL=on beat; RhythmR=Clash)

II Sonja Kohn und Herbert Otten sind bei einer Sicherheitsfirma angestellt.
Klaus hat erfahren, dass Sonja Kohn Lehrlinge überwacht.

- c Carla glaubt, dass Herbert Otten auch Kollegen überwacht.
  (ObjFoc; RhythmL=on beat; RhythmR=no Clash)
- d Carla glaubt, dass Herbert Otten auch Lehrlinge überwacht.
  (SubjFoc; RhythmL=on beat; RhythmR=Clash)

III Sonja Kohn und Herbert Ott sind bei einer Sicherheitsfirma angestellt.
Klaus hat erfahren, dass Sonja Kohn Kollegen überwacht.

- e Carla glaubt, dass Herbert Ott auch Kollegen überwacht.
  (SubjFoc; RhythmL=off beat; RhythmR=no Clash)
- f Carla glaubt, dass Herbert Ott auch Lehrlinge überwacht.
  (ObjFoc; RhythmL=off beat; RhythmR=Clash)

IV Sonja Kohn und Herbert Ott sind bei einer Sicherheitsfirma angestellt.
Klaus hat erfahren, dass Sonja Kohn Lehrlinge überwacht.

- g Carla glaubt, dass Herbert Ott auch Kollegen überwacht.
  (ObjFoc; RhythmL=off beat; RhythmR=no Clash)

h Carla glaubt, dass Herbert Ott auch Lehrlinge überwacht.
(SubjFoc; RhythmL=off beat; RhythmR=Clash)

2.2. Experiment I: Unprepared Oral Reading

The first experiment concerns the effects of linguistic rhythm and discourse context on the prosodic realization of the eight conditions (5) in spontaneous (unprepared) oral reading. Based on previous experience with this design (Kentner, 2012, 2015), we make the following assumptions: in (unprepared) oral reading, the prosodic realization reflects the interpretation assigned to the ambigous *auch*, i.e., when speakers accent *auch* (i.e., deaccent the object), they take it to associate with subject focus, otherwise with object or VP focus[2].

2.2.1. Materials, Participants, Procedure

Twenty-four item sets like (5) were developed. The items were distributed over eight lists such that items and conditions were counterbalanced across the lists with each list containing exactly one condition from each item set. Additionally, each list contained 64 filler items from four unrelated experiments and three practice items not connected to any of the experimental items, yielding a total of 91 items. With the exception of the three initial practice items, the item order was determined by pseudo-randomization (van Casteren and Davis, 2006) (for each participant individually) such that items from the same experiment had a minimal distance of two intervening items from other experiments and items from the same experimental condition were separated by at least three fillers.

Twenty-four members (19 female) of the Goethe-University community (Frankfurt, Germany) took part in the experiment. All participants were native speakers of German with normal or corrected-to-normal vision per self report. Participants were not informed about the purpose of the experiment before the experiment began; they were debriefed after the experiment ended. The age range was between 19 and 50 years old.

The experiment took place in a silent office at Goethe University in single sessions for each participant. Participants were seated in front of a 21.5-inch computer screen and equipped with a microphone head set (Shure) attached to an R-44 digital recorder.

All 91 items of each list were presented in a coherent slide show created with the standard settings of the Latex beamer package (Tantau et al., 2015). Each item was presented on two consecutive screen displays. The first display presented the two context sentences in the upper half and the first two words of the target sentence (in the case of this experiment: subject and verb of the matrix clause) in the middle of the screen (all text left-aligned). Upon pressing the enter button on the keyboard, the target sentence appeared in full (leaving the rest of the first display intact). Participants were asked to read the first display (i.e., the context) silently before moving on to the second display screen. To ensure spontaneous, unprepared oral reading and minimal look-ahead, participants were instructed to read out the target sentence immediately as it appeared on screen and to do so as fluently as possible. The participants were discouraged from making corrections during or after reading and to move on to the next item after reading by another button press. The productions of the participants were recorded on a digital memory card.

2.2.2. Results

All in all, (24 × 24=) 576 experimental sentences were recorded.

Two judges independently evaluated each target sentence. Their task was to determine by ear (i) whether the production was a fluent and flawless response to the target sentence, and (ii) where the nuclear accent was realized (on *auch* or on the object). In order to avoid influencing their judgment, the judges were not informed about the context that preceded the target sentence. Correspondingly, context-target inconsistencies with respect to the accentuation of *auch* were not coded as flaws.

Twenty-six sentences were scored as non-fluent or flawed by at least one judge. For 12 additional sentences, judges were unsure or did not agree as to where the nuclear accent was realized. These 38 sentences (6.6%) were discarded from further analysis.

Aggregating the 538 valid responses, *auch* was perceived as accented in about 24% of the cases (**Figure 1**). Note, that the full consideration of the context by the participants would imply 50% accented *auch* (all contexts inducing subject focus); correspondingly, only 67% of all trials were realized in accordance with the contextual conditions. We applied logistic mixed models (Bates et al., 2014) in the statistical computing environment R (R Core Team, 2014) to assess the effects of Context, RhythmLeft and RhythmRight as well as their interactions on the realization of accent. Accent was treated as a categorial variable (accent on *auch*: 1, accent on the object: 0). The fixed effects were coded as orthogonal sum contrasts to ensure minimal correlation; the contrast coding is shown in **Table 1**.

Random intercepts were included for participants and items. The results of the logistic mixed model are tabulated in **Table 2**. Over and above the preference for unaccented *auch*, the preceding Context significantly affected the realization of accent on *auch*. As expected, *auch* is more often accented when the preceding context renders the object given (accentuation of *auch* in 41% of cases), than when the object is new (accentuation in 8% of the cases). RhythmLeft has a weaker but still significant effect: *auch* is more likely to be accented when it falls onto the beat that is established by the rhythmic context to the left (*auch* accented in 28% of the cases) than when it is in off-beat position (20% accented).

The effect of RhythmRight on accentuation is not significant by itself but a significant three-way interaction Context:RhythmL:RhythmR attests the expected avoidance of stress clash (preference for leaving *auch* unaccented when the following syllable is stressed) in subject focus contexts when *auch*

[2]Note that we cannot know what stage of the comprehension process exactly is reflected in the prosodic form of a read utterance because the articulation necessarily follows several, but presumably not all, interpretative processes in oral reading. It is very well possible that *auch* may be realized with accent although it was initially interpreted as associating with subject focus and vice versa. We assume here that accent on *auch* implies a preponderance of subject focus interpretation during processing and, conversely, unaccented *auch* implies a predominant object focus interpretation.

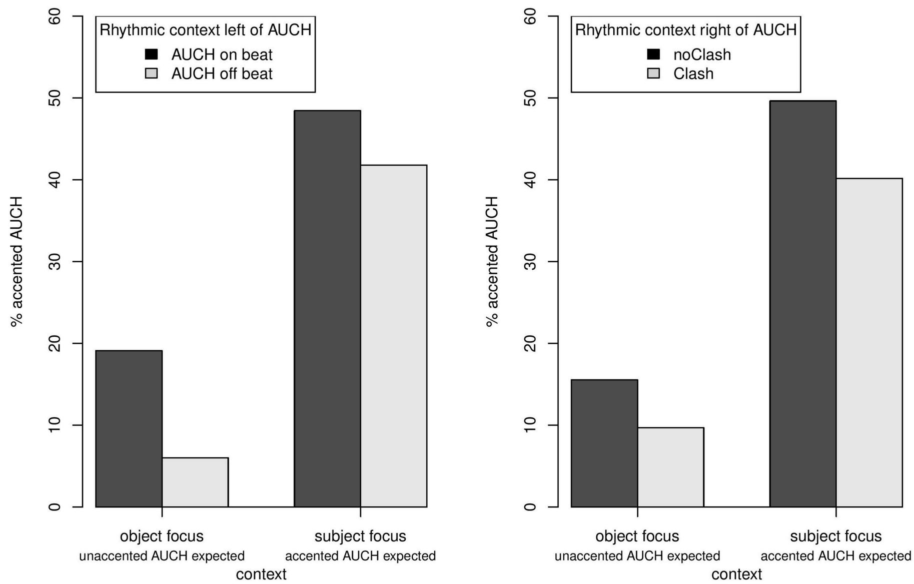


FIGURE 1 | Percentages of accented *auch* broken down by context (in both panels, the left pair of bars represent object focus, the right pair of bars represent subject focus) and rhythmic environment to the left (left panel: *auch* on beat vs. off beat) and rhythmic environment to the right (right panel: initial vs. non-initial stress on object).

TABLE 1 | The contrast coding used for the statistical analyses.

Context	
−1: accent unexpected (object focus)	1: accent expected (subject focus)
RhythmLeft	
−1: "auch" on beat	1: "auch" off beat
RhythmRight	
−1: non-initial stress on object (no clash)	1: initial stress on object (clash)

TABLE 2 | Results of logistic mixed model on perceived accentuation of *auch* in experiment I (unprepared reading).

Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)
Context	1.27759	0.15913	8.029	<0.001
RhythmL	−0.30785	0.15236	−2.021	0.0433
RhythmR	−0.08269	0.15223	−0.543	0.5870
Context:RhythmL	0.15835	0.15199	1.042	0.2975
Context:RhythmR	−0.01391	0.15203	−0.092	0.9271
RhythmL:RhythmR	0.22844	0.15232	1.500	0.1337
Context:RhythmL:RhythmR	−0.35593	0.15263	−2.332	0.0197

is in off-beat position, and in object focus contexts when *auch* is on-beat.

2.2.3. Phonetic Realization of Accented vs. Unaccented *auch*

As mentioned above, perceived accentuations of the target word *auch* are comparatively rare, i.e., in only about 24% of the cases. However, accentuation would be required in all subject focus contexts, i.e., 50% of the cases. The reason for this discrepancy is most likely due to the task (unprepared oral reading) and the general preference for function words to remain unaccented (Bader, 1996).

In order to exclude misperception by the judges as a source for this data pattern, their assessment was validated by means of a phonetic analysis. Also, since listeners may perceive prominence patterns on syllable sequences in context even in the absence of definite acoustic cues for such a pattern (Dilley and McAuley, 2008), a validation of the raters' judgments seems appropriate. Hence, the syllable durations and pitch contours of sentences with perceived accented and unaccented *mehr* were compared.

The 538 valid responses were annotated by a student assistant who was not informed about the purpose and the conditions of the experiment. Using the text-grid device in praat (Boersma and Weenink, 2010), the syllables in the critical region around *auch* were demarcated by hand, with two syllables preceding (corresponding to the subject of the embedded clause) and three syllables following *auch* (corresponding to the object). Each annotated syllable was split into three equal-sized intervals for which the mean F0 was recorded. The raw mean F0 values were normalized using the inverse of the utterance wide mean F0 multiplied by the global mean F0 (aggregated over speakers and items) as normalizing factor. The normalized values were interpolated to create an average time-normalized pitch contour. The plot in **Figure 2** juxtaposes the time normalized contours for perceived accented (black) vs. perceived unaccented versions (gray) of *auch*. Apparently, accented *auch* (as revealed by a higher F0 on that word) coincides with a higher F0 rise on the preceding subject and deaccentuation of the object. These effects are perfectly in line with the expectations: *auch* is accented when

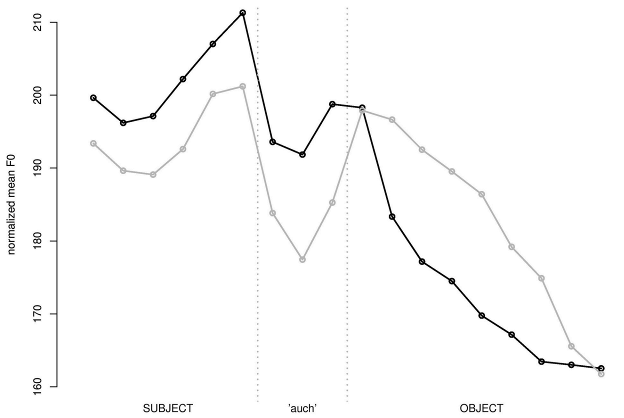


FIGURE 2 | Time normalized pitch contour (normalized mean pitch) for perceived accented (black) and unaccented (gray) versions of *auch*.

it associates with subject focus, and focus may induce the higher prominence of the subject. In the same condition, the object is given, which may explain the deaccentuation on that constituent.

A linear mixed model evaluating the effect of accentuation on the (logarithmized) duration of *auch* confirms that it is significantly longer when it is perceived as accented (raw mean duration = 224 ms) than when it is not (raw mean duration = 197 ms) (Estimate: 0.126, Std.Err: 0.028, $t = 4.47$).

2.2.4. Discussion

The oral reading experiment confirms that Context, preceding rhythm (RhythmLeft), and stress clash (RhythmRight) have (interactive) effects on the realization of accent on the ambiguous focus particle *auch*. The strong effect of Context on accentuation confirms that speakers do pay attention to the previous discourse when reading out the ambiguous target sentence. However, there is a high rate of context-target inconsistencies, especially for contexts inducing subject focus (only 40% of *auch* in subject focus conditions were perceived as accented). The high rate of inconsistencies shows that the task (unprepared oral reading) is appropriate to assess which reading is preferred in spontaneous reading without previous skimming. The clear preference for the object focus realization may be due to the fact that *auch* associating with subject focus may be expressed in a different way, i.e., with (unaccented) *auch* preceding the focused subject as in (6).

(6) Carla glaubt, dass auch Herbert Lehrlinge überwacht.
Carla thinks that Herbert, too, supervises apprentices.

Auch preceding the subject may in fact be a more natural expression of subject focus for three reasons: First, with this word order, there is no ambiguity as to the association of the focus operator—association of *auch* with the object is impossible / ungrammatical in (6). Secondly, in (6) but not in the subject focus versions of (5), the focus particle is left-adjacent to its scope domain. In this configuration, the focus particle acts as a herald for the focus domain—in contrast, postponed *auch* requires retrospective confirmation of the focus domain or, worse, reanalysis. The third reason is prosodic in nature: postponed *auch* associating with subject focus bears an accent (Féry, 2009); accent on function words, however, are highly marked[3]. Be this as it may, postponed *auch* is perfectly acceptable and grammatical in the subject focus contexts in (5).

The significant main effect of RhythmLeft confirms the hypothesis that the preceding trochaic beat—as established by the sequence of lexical prominences—leads to rhythmic expectations concerning upcoming material. As predicted, readers are less likely to accent *auch* when it falls onto an off-beat position or more likely to accent *auch* when it is in a strong position of the beat.

In contrast to our previous experiments (Kentner, 2012, 2015; McCurdy et al., 2013), readers did not systematically avoid accentuation in the context of a potential stress clash to the right of *auch*—the effect of RhythmRight remains non-significant by itself. Rather, the significant three-way interaction shows that the effectiveness of the RhythmRight manipulation depends on the disposition of both RhythmLeft and Context. We will return to the lack of this effect in the General Discussion.

Under the assumption that the prosodic realization of *auch* reflects the readers' interpretation of the focus particle, we may submit that all three factors contribute to the way in which speakers interpret the target sentence.

However, experiment I does not allow firm conclusions to be drawn about the interplay of prosodic rhythm and contextual information in reading comprehension. So far, we have only evaluated data pertaining to speech production, which is known to lag behind interpretative processes in oral reading (Levin and Addis, 1979; Inhoff et al., 2011; Laubrock and Kliegl, 2015).

The eye-movement record provides data that is certainly more time-sensitive and thus more informative about the impact of implicit prosody and context in sentence comprehension (Clifton et al., 2007; Vasishth et al., 2013).

2.3. Experiment II: Silent Reading

Experiment 2 was an eyetracking version of Experiment 1.

2.3.1. Methods

2.3.1.1. Materials

The 24 item sets from Experiment 1 were again distributed over eight lists with items and conditions counterbalanced across the lists. Each list contained exactly one condition from each item set. In addition, 60 items from two unrelated experiments were interspersed as fillers. Each list was preceded by five practice items, yielding a total of 89 items per participant.

2.3.1.2. Participants

Fifty-two native speakers of German from the Berlin area participated in the experiment for partial course credit or for payment. All participants reported normal or corrected-to-normal vision.

[3]In this context, it is interesting to note that the overall percentage of accents realized on *auch* in this experiment (25%) exactly corresponds with the percentage of perceived accented versions of the critical word *mehr* in the unprepared oral reading experiment in Kentner (2012), cf. example in (7).

2.3.1.3. Apparatus and Procedure

Fixation time measures were gathered from the participants' right eye using an SMI (SensoMotoric Instruments) IView-X eye-tracker running at a sampling rate of 240 Hz (0.025 degree tracking resolution, and <0.5 degree gaze position accuracy). A chin rest was used to ensure stability. The chin rest was placed 55 cm from a 17 inch monitor (1024 × 768 pixel resolution). The angle per character was 0.3 degrees (3.8 characters per degree of visual angle). Stimulus presentation was controlled by Presentation software. Eye-gaze calibration was carried out at the beginning of the experiment, and calibration quality was monitored by the experimenter, with recalibration every ten trials, or more frequently if necessary.

Before each trial, the participant fixated upon a black dot in the center of the left side of the screen to ensure calibration quality. Successful fixation of the dot triggered the appearance of the context sentence, at which point the participant read it through and pressed a continuation button. The fixation point appeared once more at the same location, and after one second the point was replaced by the target sentence. Fixation data were gathered continuously throughout each trial. When the participant finished reading the sentence, either he or she was required to answer a yes/no comprehension question in the case of 36 out of the 60 filler items, or a fixation cross appeared on screen announcing the next trial. The two context sentences were broken into two lines, target sentences always appeared on one line. Participants took about 45 min to complete the experiment.

2.3.2. Data Analysis

The *em* package by Logačev and Vasishth (2006) was used to calculate the dependent measures from the raw output of the eye-tracking software. Inspection of the individual eye movement patterns revealed mis-calibration in three participants. Data from these participants were discarded, i.e., we considered data from 49 subjects.

2.3.2.1. Regions of interest

We analyzed the eye movement data from four consecutive words, starting in the word preceding the focus particle *auch* up to the end of the sentence. The word preceding *auch* is the subject of the embedded clause and the locus of the RhythmLeft manipulation—if this word is disyllabic (trochaic), *auch* falls on a strong position of the the beat established by the preceding trochaic rhythm; conversely, *auch* is off beat relative to the established rhythm (i.e., on a weak position) in the case of a monosyllabic subject. At the same time, the subject is a potential bearer of the focus *auch* associates with. The second word of interest is the ambiguously attachable *auch*. The following object or, more precisely, the givenness or newness of the object, determines the interpretation of *auch*. When the object was already mentioned in the context, *auch* necessarily associates with subject focus and would bear nuclear accent in a spoken rendition of the sentence. In this case, the object would be de-accented. If the object was not previously mentioned, *auch* associates with object focus, with the object bearing the main sentence accent in a spoken rendition. The disambiguating object is also the locus of the RhythmRight manipulation. If starting in a stressed syllable, there is a potential stress clash if the critical *auch* would bear an (implicit) accent, which would be the case when *auch* is interpreted as associating with subject focus. The last word of the sentence is the verb of the subordinate clause. Irrespective of the experimental condition, this word is given in the discourse context.

2.3.2.2. Reading measures

For the four regions of interest, we report three kinds of word reading times that were extracted from the eye-tracking data:

- First-pass reading time (FPRT, a.k.a. gaze duration): the summed duration of all fixations on a word before a fixation on any other word—given that no word to the right of the current word was fixated.
- Regression path duration (RPD) or Go-past time: summed duration of all fixations from the first fixation on the current word up to (but not including) the first fixation on a word further to the right. Note that this includes regressive fixations on words to the left of the current word.
- Total reading time (TFT): Summed duration of all fixations on a word

Fixations shorter than 50 ms were removed and treated as missing values. Fixation durations were log-transformed for inferential statistics.

2.3.3. Results

Mean reading times and standard errors for the eight conditions are tabulated in **Table 3**. Linear mixed effects models (LMM, Bates et al., 2014) were employed to assess the influence of the fixed factors Context, RhythmLeft and RhythmRight, and their interactions on reading times. To ensure minimal correlations among the fixed effects, they were coded as orthogonal sum contrasts. Only the intercepts for participants and items were included as random effects. LMMs with a "maximal" random effect structure with varying intercepts and slopes, as advocated in Barr et al. (2013), did not converge or led to pathological estimates of the random effect correlations.

In order to ensure that the results of the models with parsimonious random effect structure hold, we also fit Bayesian LMMs with the maximal random effect structure justified by the design of the experiment. In contrast to conventional LMMs, Bayesian LMMs always allow for complex random effect structures because the weakly informative priors used in the modeling will ensure that the posterior distribution of a parameter will be centered around 0 if there is not enough data to estimate the true value of the parameter. Since the results of the Bayesian analysis largely conform to the outcome of the conventional analysis, we report only the latter. The details of the Bayesian LMM are given in the Appendix.

Figure 3 shows raw first pass reading times (upper row), regression path durations (middle row) and total fixation times (lower row) in milliseconds for three consecutive words starting in the ambiguously attachable *auch* through to the sentence final verb. The reading times are broken down by the factors Context (dark bars = Object focus, light bars = Subject focus) and RhythmLeft (*auch* off beat vs on beat). The factor RhythmRight is disregarded for reasons of clarity. The disambiguating object

TABLE 3 | Raw reading times (FPRT, RPD, and TFT) with standard errors, broken down by condition, for the four critical regions (Subject of the embedded clause, focus particle *auch*, Object, and sentence final Verb.

Context	RhythmL	RhythmR	FPRT (SE)			
			Subject	*auch*	Object	Verb
Object foc	On beat	Unstressed	274 (19)	242 (16)	292 (21)	302 (26)
		Stressed	295 (22)	248 (16)	317 (26)	351 (29)
	Off beat	Unstressed	252 (15)	246 (16)	275 (19)	337 (31)
		Stressed	253 (15)	227 (15)	280 (23)	331 (28)
Subject foc	On beat	Unstressed	271 (18)	250 (16)	277 (20)	313 (24)
		Stressed	288 (24)	241 (15)	270 (19)	316 (33)
	Off beat	Unstressed	272 (20)	258 (15)	279 (19)	273 (22)
		Stressed	279 (19)	262 (16)	266 (21)	303 (25)
			RPD (SE)			
Object foc	On beat	Unstressed	317 (25)	279 (28)	419 (61)	692 (79)
		Stressed	333 (31)	313 (31)	430 (53)	654 (79)
	Off beat	Unstressed	302 (35)	268 (18)	365 (36)	672 (79)
		Stressed	310 (29)	266 (26)	408 (45)	649 (89)
Subject foc	On beat	Unstressed	339 (36)	265 (18)	387 (44)	587 (65)
		Stressed	354 (32)	283 (21)	367 (37)	623 (98)
	Off beat	Unstressed	323 (33)	282 (20)	389 (44)	644 (90)
		Stressed	318 (28)	322 (31)	388 (53)	659 (96)
			TFT (SE)			
Object foc	On beat	Unstressed	370 (29)	321 (33)	383 (34)	361 (28)
		Stressed	405 (40)	302 (22)	384 (36)	388 (35)
	Off beat	Unstressed	318 (29)	291 (21)	343 (29)	396 (35)
		Stressed	326 (24)	268 (21)	380 (37)	383 (29)
subject foc	On beat	Unstressed	365 (29)	285 (20)	338 (26)	344 (26)
		Stressed	378 (33)	282 (19)	336 (29)	353 (35)
	Off beat	Unstressed	330 (28)	314 (24)	353 (34)	355 (33)
		Stressed	325 (23)	313 (26)	324 (26)	391 (42)

(middle column) determines the attachment site of *auch* and hence disambiguates the sentence.

We report inferential statistics on the reading data for the subject preceding *auch*, i.e., the locus of RhythmLeft manipulation, and the three following words.

2.3.3.1. Subject preceding auch

All three reading measures reveal significant main effects of the RhythmLeft manipulation on the subject preceding *auch* (t-values$\geq|2|$) with longer reading times for disyllabic (trochaic) subjects compared to monosyllabic ones. All other main effects and interactions are non-significant with t-values$\leq|1.2|$, cf. **Table 4**.

2.3.3.2. Ambiguous auch

The eye-tracking data show reading times on *auch* to be affected by the factors Context and RhythmLeft (**Table 5**): FPRTs on *auch* are significantly longer when the Context manipulation requires subject focus.

In addition, RPDs and TFTs on *auch* reveal a significant interaction of Context and RhythmLeft to the effect that reading times on this word are longer when the position of the monosyllable *auch* with respect to the beat established by the preceding rhythmic context is in conflict with the contextually determined (implicit) accentuation or de-accentuation of *auch*. All other main effects and interactions on this word remain non-significant.

2.3.3.3. Object following auch and sentence final verb

Reading times on the object (**Table 6**) reveal a main effect of context. FPRTs and TFTs are significantly longer in the object focus condition, i.e., for objects that were not mentioned in the previous context. Note that in the subject focus

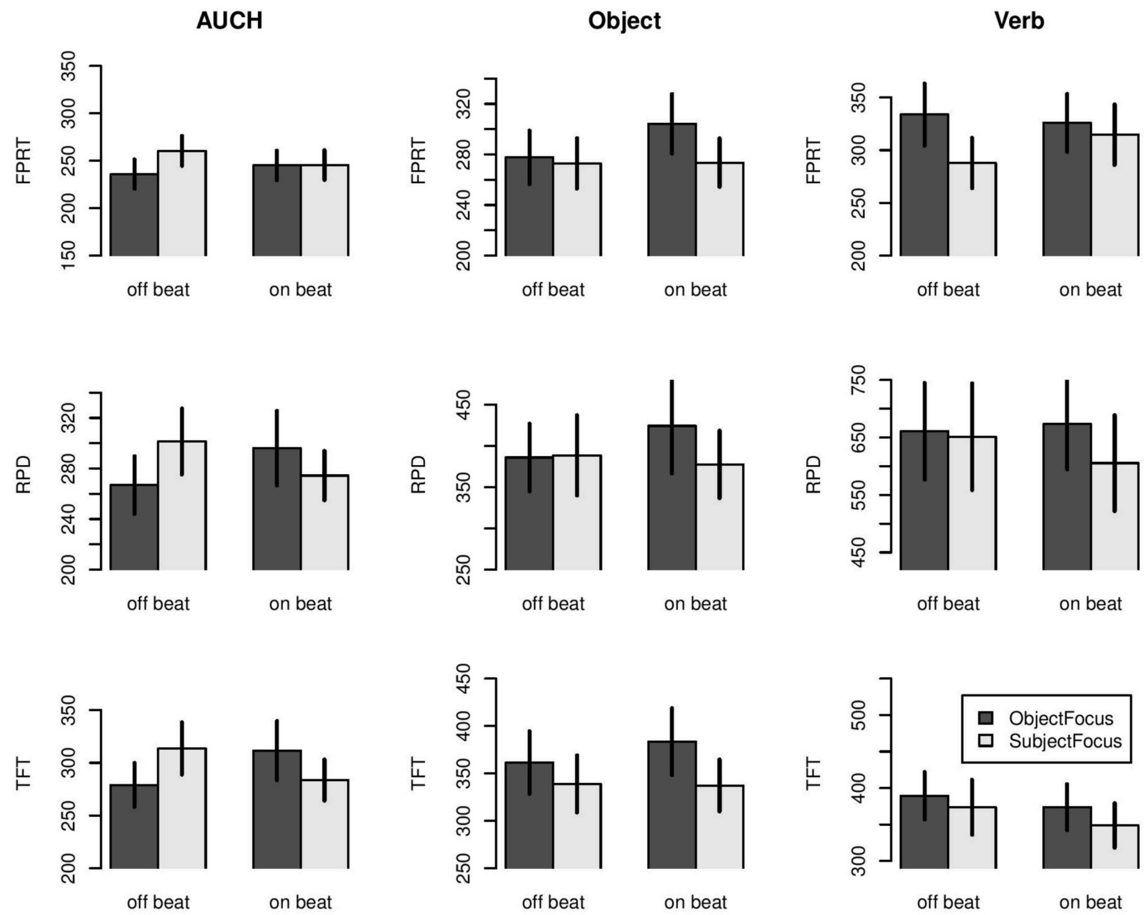


FIGURE 3 | First pass reading times (upper row), regression path durations (middle row) and total fixation times (lower row) for three consecutive words starting in the ambiguously attachable *auch* through to the sentence final verb. Reading times are broken down by the factors Context (dark bars = Object focus, light bars = Subject focus) and RhythmLeft (*auch* off beat vs on beat). The factor RhythmRight is disregarded here. Error bars correspond to one standard error. Note that the contextual givenness of the Object determines the association of *auch*.

TABLE 4 | Results of LMM for reading times on the Subject.

Subject	FPRT		RPD		TFT	
	Est.(SE)	*t*	Est.(SE)	*t*	Est.(SE)	*t*
Context	0.007(0.01)	0.55	0.017(0.02)	1.06	−0.000(0.02)	−0.02
RhythmL	−0.026(0.01)	−2.00+	−0.034(0.02)	−2.2*	−0.071(0.02)	−4.46*
RhythmR	0.015(0.01)	1.13	0.016(0.02)	1.03	0.013(0.02)	0.85
Con×RhL	0.015(0.01)	1.20	0.002(0.02)	0.15	0.005(0.02)	0.35
Con×RhR	0.001(0.01)	0.04	0.002(0.02)	0.13	−0.007(0.02)	−0.46
RhL×RhR	−0.01(0.01)	−0.8	−0.006(0.02)	−0.38	0.001(0.02)	0.08
C×RL×RR	0.007(0.01)	0.55	−0.008(0.02)	−0.49	−0.003(0.02)	−0.2

Statistics significant at the level of $\alpha = 0.05$ (i.e., t-values≥|2|) are highlighted by an asterisk, or, if the 95% credible interval of the corresponding parameter in the Bayesian LMM includes 0, by a + −sign.

condition, participants already read the same object-verb sequence in the context. The same effects of context were found for FPRTs and TFTs on the sentence final verb (**Table 7**). In addition, there is a significant three-way interaction in FPRTs on the sentence-final verb, revealing a late influence of RhythmRight modulating effects of Context and RhythmLeft. RPDs on the object and the verb remain largely uninformative.

TABLE 5 | Results of LMM for reading times on the focus particle *auch*.

auch	FPRT		RPD		TFT	
	Est.(SE)	***t***	**Est.(SE)**	***t***	**Est.(SE)**	***t***
Context	0.27(0.01)	2.14+	0.019(0.015)	1.27	0.012(0.02)	0.79
RhythmL	0.008(0.01)	0.67	−0.003(0.015)	−0.19	0.002(0.02)	0.19
RhythmR	−0.013(0.01)	−1.04	0.022(0.015)	1.48	−0.015(0.02)	−0.97
Con×RhL	0.025(0.01)	1.97	0.038(0.015)	2.53*	0.041(0.02)	2.66*
Con×RhR	0.006(0.01)	0.48	0.012(0.015)	0.79	0.011(0.02)	0.74
RhL×RhR	−0.008(0.01)	−0.62	−0.017(0.015)	−1.17	−0.014(0.02)	−0.9
C×RL×RR	0.016(0.01)	1.29	0.022(0.015)	1.48	0.004(0.02)	0.25

Statistics significant at the level of $\alpha = 0.05$ (i.e., t-values≥|2|) are highlighted by an asterisk, or, if the 95% credible interval of the corresponding parameter in the Bayesian LMM includes 0, by a $^{+}$-sign.

TABLE 6 | Results of LMM for reading times on the Object.

Object	FPRT		RPD		TFT	
	Est.(SE)	***t***	**Est.(SE)**	***t***	**Est.(SE)**	***t***
Context	−0.397(0.19)	−2.08+	−0.026(0.02)	−1.43	−0.046(0.02)	−2.9*
RhythmL	−0.345(0.19)	−1.8	−0.02(0.02)	−1.09	−0.016(0.02)	−1.00
RhythmR	−0.054(0.19)	−0.01	0.002(0.02)	0.13	−0.000(0.02)	−0.02
Con×RhL	0.327(0.19)	1.71	0.016(.02)	0.91	0.016(0.02)	1.00
Con×RhR	−0.28(0.19)	−1.46	−0.028(0.02)	−1.56	−0.016(0.02)	−0.99
RhL×RhR	−0.165(0.19)	−0.86	0.003(0.02)	0.17	0.006(0.02)	0.34
C×RL×RR	0.024(0.19)	0.32	−0.006(0.02)	30.31	−0.011(0.02)	−0.7

Statistics significant at the level of $\alpha = 0.05$ (i.e., t-values≥|2|) are highlighted by an asterisk, or, if the 95% credible interval of the corresponding parameter in the Bayesian LMM includes 0, by a $^{+}$-sign.

TABLE 7 | Results of LMM for reading times on the Verb.

Verb	FPRT		RPD		TFT	
	Est.(SE)	***t***	**Est.(SE)**	***t***	**Est.(SE)**	*t*
Context	−0.041(0.02)	−2.38+	−0.045(0.02)	−1.87	−0.037(0.02)	−2.12+
RhythmL	−0.022(0.02)	−1.29	−0.001(0.02)	−0.05	0.01(0.02)	0.55
RhythmR	0.026(0.02)	1.51	−0.005(0.02)	−0.19	0.014(0.02)	0.77
Con×RhL	−0.014(0.02)	−0.79	0.017(0.02)	0.71	0.004(0.02)	0.24
Con×RhR	−0.011(0.02)	−0.63	0.018(0.02)	0.74	0.004(0.02)	0.22
RhL×RhR	−0.002(0.02)	−0.11	0.01(0.02)	0.39	0.001(0.02)	0.04
C×RL×RR	0.038(0.02)	2.2*	0.032(0.02)	1.33	0.022(0.02)	1.25

Statistics significant at the level of $\alpha = 0.05$ (i.e., t-values≥|2|) are highlighted by an asterisk, or, if the 95% credible interval of the corresponding parameter in the Bayesian LMM includes 0, by a $^{+}$-sign.

2.4. Discussion

The data reveal an immediate if transient interaction of RhythmLeft and Context on *auch*, suggesting that readers consult and consider both sources of information simultaneously while forming an interpretation for the ambiguously attachable word. The eye tracking record attests enhanced reading effort if contextual and prosodic constraints on the interpretation of *auch* are in conflict. More concretely, reading times increase significantly when the monosyllable *auch* needs to be accented (subject focus) but falls onto a weak (off-beat), and hence less accentable, position with respect to the established rhythm.

In addition, the FPRTs show an early main effect of Context on *auch*. Note that both the main effect of Context and the interaction are in force before the disambiguating object has been fixated (as revealed by FPRT and RPD on this word). Since the contextual manipulation hinges on the givenness of the object directly following *auch*, the main effect in FPRT and the interactions of RhythmLeft and Context in FPRT and RPD suggest parafoveal preview of the object—given the shortness of *auch* this is a likely scenario.

Apart from the early Context×RhythmLeft interaction, the main effect of Context varies in polarity throughout the critical regions. On *auch*, the data suggest that readers experience more difficulty with subject focus readings. As noted above in Section 2.2.4, we assume that postponed *auch* in the subject focus reading is relatively marked and may therefore be the dispreferred reading. In contrast, on both the object and the verb, reading times are significantly shorter in the case of subject focus contexts. The reason for this disparity is likely due to the contextual givenness of the object in the subject focus conditions: in general, readers make shorter fixations on words that they have encountered shortly before. This familiarity advantage apparently overrides any effect stemming from the syntactic markedness of the subject focus condition.

The effect of RhythmRight is less systematic, and only becomes apparent in FPRTs on the sentence-final verb in the form of a three-way interaction. We discuss possible reasons for the weak influence of RhythmRight in the General Discussion.

3. GENERAL DISCUSSION

The two experiments reported here were designed to test the interaction of local phonological and more global, discourse-contextual information during the interpretation of structurally ambiguous sentences in oral and silent reading. In the first experiment (unprepared oral reading), we found a clear preference for the prosodic realization of the object focus reading with unaccented *auch*, a strong effect of Context (more accentuations of *auch* when the Context required the subject focus reading) and a weaker but systematic effect of RhythmLeft such that accentuation of *auch* was avoided in off-beat position. The effect of RhythmRight (avoidance of stress clash) turned out to be less systematic.

Similarly, the silent reading experiment yields an effect of Context, with reading times on the ambiguous word *auch* increased when the Context requires subject focus—we take this to confirm the general preference for the object focus reading that we found in the oral reading experiment. Moreover, a significant Context×RhythmLeft interaction on *auch* confirms that global discourse context and local prosodic rhythm conspire to condition the way the sentence is being interpreted. These effects were detected in so-called early reading time measures, which could suggest that they reflect early stages in the comprehension process (Clifton et al., 2007). Importantly, the Context×RhythmLeft interaction on *auch* emerges before readers fixated the disambiguating object. Therefore, the effect is unlikely to be driven by reanalysis processes; rather, it points to a guiding function of implicit rhythm in parsing, in line with findings by Breen and Clifton (2013) and Kentner (2012). Given the early influence of prosody on syntactic parsing, the present results are difficult to reconcile with accounts like the ones by Augurzky (2006), Kondo and Mazuka (1996), or Koriat et al. (2002), all of which consider syntactic structure building to be a prerequisite for the prosodic analysis in reading (see Kentner, unpublished, for a similar point).

What, then, is the nature of the early RhythmLeft×Context interaction affecting reading times on *auch*? Contextual givenness and low level linguistic rhythm are, at first sight, independent phenomena; an interaction may therefore seem surprising. While contextual givenness affects, even determines, the eventual association of the ambiguous focus particle *auch*, there is no obvious reason why the RhythmLeft manipulation—i.e., the variation of the syllabic structure of the subject preceding *auch* and, hence, the continuation of the established beat—should condition the interpretation of *auch*. However, the link becomes explicable when considering the prosodic consequences of the contextually determined interpretation of *auch*. Recall that a contextually given object induces *auch* to be associated with subject focus. In this case, *auch* bears the main sentence accent in a spoken rendition, and a corresponding implicit accent in silent reading. Conversely, if *auch* associates with object focus, it remains unaccented and the main accent is realized on the (newly introduced) object. The RhythmLeft manipulation engenders prosodic constellations that either facilitate or hinder accentuation of *auch*: If *auch* is "on beat" relative to the preceding trochaic rhythm (as established by the successive alternation of lexically stressed and unstressed syllables), accentuation is considered easy but it is considered hard when *auch* is "off beat." The Context×RhythmLeft interaction reflect this: When, in order to comply with a contextual imperative, accentuation of *auch* is required but, at the same time, accentuation is hard on rhythmic grounds, readers tend to avoid accentuation in oral reading (Experiment I) or—in the case of silent reading (Experiment II)—the computation of the required structure is effortful and reading times increase. The results are therefore consistent with the early involvement of implicit prosodic rhythm and accentuation in written sentence comprehension. This interpretation is generally in line with our previous findings on the role of implicit prosody and rhythm in reading. However, there are notable differences concerning the details of the rhythmic and contextual effects, which we discuss below.

Kentner (2012) and McCurdy et al. (2013) explored the influence of linguistic rhythm on the interpretation of syntactically ambiguous structures like (7) in which the requirement for accentuation of the ambiguous word sequence *nicht mehr* depended upon its syntactic status as either a temporal adverb [(7-a), requiring unaccented *mehr*] or a negated comparative quantifier [(7-b), main phrase accent on *mehr*]. The rhythmic manipulation targeted the word following *nicht mehr*, featuring three-syllabic verbs with either initial or non-initial lexical stress.

(7) Tim meint, dass man...
Tim thinks that one...

a. nicht mehr {nachweisen / ermitteln} kann, wer der
not more {determine / find out} can, who the
Täter war.
culprit was.

b. nicht MEHR {nachweisen / ermitteln} kann, als
not more {determine / find out} can, than
die Zeit.
the time.

Both studies showed increased reading times for structures that engendered a stress clash (accented comparative *mehr* followed by a verb with initial stress) compared to non-clashing conditions. Kentner (2012) also reports an oral reading study in which readers avoid accentuation of the critical word when this leads to rhythmically infelicitous stress clash. While in Kentner's (2012) silent reading study, the effect of stress clash was detected at the disambiguating region following the verb, McCurdy et al. suggest that the rhythmic factor already affected the ambiguous word *mehr* itself. The rhythmic factor in those studies corresponds to the RhythmRight manipulation in the present experiments: i.e., the variation concerns the position of lexical stress on the word following a syntactically ambiguous, variably accentable word with the potential consequence of stress clash if the following word bears initial stress. However, comparing the results at hand with those by Kentner (2012) and McCurdy et al. (2013), it becomes clear that the present effect of RhythmRight on reading behavior deviates from the previous effects in that it is very limited and is detectable only relatively late.

A conceivable explanation for this disparity lies in the difference of the linguistic structures under study. As pointed out above, the structures of the present experiment (with *auch*) are superficially similar to the previously used items (with *nicht mehr*) in that the rhythm-syntax manipulation is brought about by a variably accentable, syntactically ambiguous word followed by a word featuring either initial or non-initial lexical stress. Despite this similarity, however, the syntactic relation of the ambiguous word with the following word differs between the experiments and experimental conditions: Consider first the case of the accented comparative quantifier *mehr* in (8-a): This word fills the object position of the following verb, and is thus part of the verb phrase, which, under standard assumptions, is mapped onto a prosodic phrase (Truckenbrodt, 2006). In contrast, accented *auch* in (8-b) is associated with the preceding subject and thus syntactically disjoint from the following object. A phrase boundary separating the two prominent syllables is a likely reason for the relatively limited effect of RhythmRight in this experiment—the boundary serves as a cesura that makes any effect of stress clash disappear (cf. Hayes, 1989; Sandalo and Truckenbrodt, 2003).

(8) a. ... [nicht MEHR nachweisen kann $]_{VP}$ als ...
... [not more determine can $]_{VP}$ than ...

b. ... [Hans AUCH $]_{NP}$ [Lehrlinge überwacht $]_{VP}$
... [Hans, too, $]_{NP}$ [apprentices supervises $]_{VP}$

Another difference between the present study and the experiment by McCurdy et al. (2013) concerns the effect of context and its interaction with the rhythmic manipulation. McCurdy et al. (2013) found only late effects of context and little interaction of context with prosodic rhythm. McCurdy et al. (2013) used a contextual priming strategy to bias the reader toward either the comparative or the temporal reading of ambiguous *nicht mehr*. There was, however, no compelling relation between the contextual bias and the resolution of the ambiguity in the target sentence. As opposed to such a loose relation between the context and the target ambiguity, the contextual manipulation of the present experiment is decisive for the correct interpretation of the ambiguous word—it hinges on the contextual givenness of the object. It may be the more compelling nature of the context sentence that led readers to take more careful note of its information when parsing the target sentence, resulting in earlier and stronger effects of context. A recent study by Logačev and Vasishth (2015) supports this view: building on work by Swets et al. (2008), Logačev and Vasishth (2015) show that contexts that are especially relevant for the interpretation of the target ambiguity may have important consequences for comprehension strategies as regards the target sentence. Specifically, they explored the nature of the ambiguity advantage that had been reported for globally ambiguous sentences (van Gompel et al., 2001, 2005), i.e., the fact that globally ambiguous sentences are read faster than non-ambiguous analogues. Logačev and Vasishth (2015) found that the presence of the ambiguity advantage depends on the nature of the comprehension questions readers were required to answer. While readers who had to answer superficial comprehension questions did show the ambiguity advantage, participants who had to respond to comprehension questions which required deeper text comprehension showed an ambiguity disadvantage, i.e., they read ambiguous sentences slower than the non-ambiguous counterparts. In the present eyetracking experiment, since readers were confronted with context sentences that are crucial for the appropriate interpretation of the target sentence, the context effect may be stronger and may have shown up early enough to directly interact with local prosodic information.

4. CONCLUSION

The two experiments presented in this study provide evidence suggesting that, during oral and silent reading, readers deploy both higher-level context and the rhythmic structure of German to disambiguate the attachment of the focus particle *auch*. A conflict between the disambiguation provided by context vs. rhythmic structure leads to a greater reading difficulty. We argue that such a conflict arises because both the contextual information, which co-determines the focus structure, as well as the prosodic rhythm, which establishes a prosodic prominence profile, affect the (implicit) accentuation of the text. This work therefore provides independent support for the claim that silent prosody plays an important role in parsing decisions, and that multiple sources of information are simultaneously deployed in resolving ambiguities.

AUTHOR CONTRIBUTIONS

GK conceived Experiments 1 and 2, conducted Experiment 1, analyzed data of Experiments 1 and 2, wrote the manuscript. SV provided the eye-tracking infrastructure, analyzed the data of Experiments 1 and 2, wrote the manuscript.

ACKNOWLEDGMENTS

We thank Kate McCurdy who recruited the participants of, and managed, the eye-tracking experiment. We are very grateful to Manfred Krifka, Director of the Zentrum für Allgemeine Sprachwissenschaft, for allowing us to use his laboratory in Berlin for the eye-tracking experiment. Publication of this research was funded by grant to GK (KE 1985/2-1) from the German Research Foundation (DFG).

REFERENCES

Altmann, G., and Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition* 30, 191–238. doi: 10.1016/0010-0277(88)90020-0

Altmann, H. (1976). *Die Gradpartikeln im Deutschen: Untersuchungen zu ihrer Syntax, Semantik und Pragmatik*. Berlin: Walter de Gruyter.

Ashby, J., and Clifton, C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition* 96, B89–B100. doi: 10.1016/j.cognition.2004.12.006

Ashby, J., and Martin, A. (2008). Prosodic phonological representations early in visual word recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 224–236. doi: 10.1037/0096-1523.34.1.224

Augurzky, P. (2006). *Attaching Relative Clauses in German: The Role of Implicit and Explicit Prosody in Sentence Processing*. Ph. D thesis, MPI for Human Cognitive and Brain Sciences, Leipzig.

Bader, M. (1996). *Sprachverstehen: Syntax und Prosodie beim Lesen*. Opladen: Westdeutscher Verlag.

Bader, M. (1998). "Prosodic influences on reading syntactically ambiguous sentences," in *Reanalysis in Sentence Processing*, eds J. Fodor, and F. Ferreira (Dordrecht: Kluwer), 1–46. doi: 10.1007/978-94-015-9070-9_1

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). *lme4: Linear Mixed-Effects Models using Eigen and S4 [Computer Program]*. R package version 1.1-7. Available online at: http://CRAN.R-project.org/package=lme4

Binder, K. S., Duffy, S. A., and Rayner, K. (2001). The effects of thematic fit and discourse context on syntactic ambiguity resolution. *J. Mem. Lang.* 44, 297–324. doi: 10.1006/jmla.2000.2754

Boersma, P., and Weenink, D. (2010). *Praat: Doing Phonetics by Computer [Computer Program]*. Version 5.1. Available online at: http://www.praat.org/ (Accessed July 2, 2010).

Breen, M. (2014). Empirical investigations of the role of implicit prosody in sentence processing. *Lang. Linguist. Compass* 8, 37–50. doi: 10.1111/lnc3.12061

Breen, M., and Clifton, C. (2011). Stress matters: effects of anticipated lexical stress on silent reading. *J. Mem. Lang.* 64, 153–170. doi: 10.1016/j.jml.2010.11.001

Breen, M., and Clifton, C. (2013). Stress matters revisited: a boundary change experiment. *Q. J. Exp. Psychol.* 66, 1896–1909. doi: 10.1080/17470218.2013.766899

Chafe, W. (1988). Punctuation and the prosody of written language. *Writ. Commun.* 5, 395–426. doi: 10.1177/0741088388005004001

Clifton, C., Staub, A., and Rayner, K. (2007). "Eye movements in reading words and sentences," in *Eye Movements: A Window on Mind and Brain*, eds R. van Gompel, M. Fischer, W. Murray, and R. L. Hill (Amsterdam: Elsevier), 341–372.

Dilley, L., and McAuley, J. (2008). Distal prosodic context affects word segmentation and lexical processing. *J. Mem. Lang.* 59, 294–311. doi: 10.1016/j.jml.2008.06.006

Drubig, H.-B. (1994). *Island Constraints and the Syntactic Nature of Focus and Association with Focus*. Arbeitsbericht (SFB 340), University of Tuebingen.

Féry, C. (2009). Postponed auch: where does its accent come from? *Snippets* 20, 23–27. Available online at: http://www.ledonline.it/snippets/allegati/snippets20000.pdf

Féry, C., and Samek-Lodovici, V. (2006). Focus projection and prosodic prominence in nested foci. *Language* 82, 131–150. doi: 10.1353/lan.2006.0031

Fodor, J. D. (2002). "Psycholinguistics cannot escape prosody," in *Proceedings of the 1st International Conference on Speech Prosody* (Aix-en-Provence).

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychol. Bull.* 123, 71–99. doi: 10.1037/0033-2909.123.1.71

Gabry, J., and Goodrich, B. (2016). *Rstanarm: Bayesian Applied Regression Modeling via Stan [Computer Program]*. R package version 2.9.0-1. Available online at: https://cran.rstudio.com/web/packages/rstanarm/

Gussenhoven, C. (1983). Focus, mode and the nucleus. *J. Linguist.* 19, 377–417. doi: 10.1017/S0022226700007799

Hayes, B. (1989). "The prosodic hierarchy in meter," in *Rhythm and Meter*, eds P. Kiparsky, and G. Youmans (Orlando, Academic Press), 201–260.

Hirose, Y. (2003). Recycling prosodic boundaries. *J. Psycholinguist. Res.* 32, 167–195. doi: 10.1023/A:1022448308035

Hwang, H., and Steinhauer, K. (2011). Phrase length matters: the interplay between implicit prosody and syntax in Korean garden path sentences. *J. Cogn. Neurosci.* 23, 3555–3575. doi: 10.1162/jocn_a_00001

Inhoff, A. W., Solomon, M., Radach, R., and Seymour, B. A. (2011). Temporal dynamics of the eye–voice span and eye movement control during oral reading. *J. Cogn. Psychol.* 23, 543–558. doi: 10.1080/20445911.2011.546782

Jacobs, J. (1983). *Fokus und Skalen: Zur Syntax und Semantik der Gradpartikeln im Deutschen*. Tübingen: Max Niemeyer.

Jun, S. (2003). Prosodic phrasing and attachment preferences. *J. Psycholinguist. Res.* 32, 219–249. doi: 10.1023/A:1022452408944

Kentner, G. (2012). Linguistic rhythm guides parsing decisions in written sentence comprehension. *Cognition* 123, 1–20. doi: 10.1016/j.cognition.2011.11.012

Kentner, G. (2015). "Stress clash hampers processing of noncanonical structures in reading," in *Rhythm in Cognition and Grammar: A Germanic Perspective*, eds R. Vogel, and R. van de Vijver (de Gruyter, Berlin), 111–135. doi: 10.1515/9783110378092.111

Kondo, T., and Mazuka, R. (1996). Prosodic planning while reading aloud: On-line examination of Japanese sentences. *J. Psycholinguist. Res.* 25, 357–381.

Koriat, A., Greenberg, S. N., and Kreiner, H. (2002). The extraction of structure during reading: evidence from reading prosody. *Mem. Cogn.* 30, 270–280. doi: 10.3758/BF03195288

Krifka, M. (2006). "Can focus accenting be eliminated in favor of deaccenting given constituents?," in *Proceedings of the Ninth Symposium on Logic and Language* (Besenyötelek), 107–119.

Ladd, D. R. (1996). *Intonational Phonology*. Cambridge: CUP.

Laubrock, J., and Kliegl, R. (2015). The eye-voice span during reading aloud. *Front. Psychol.* 6:1432. doi: 10.3389/fpsyg.2015.01432

Levin, H., and Addis, A. (1979). *The Eye-Voice Span*. Cambridge: MIT Press.

Logačev, P., and Vasishth, S. (2006). *The em Package for Computing Eyetracking Measures*. Potsdam: University of Potsdam.

Logačev, P., and Vasishth, S. (2015). A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. *Cogn. Sci.* doi: 10.1111/cogs.12228. [Epub ahead of print].

MacDonald, M. C., Perlmutter, N. J., and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676–703. doi: 10.1037/0033-295X.101.4.676

McCurdy, K., Kentner, G., and Vasishth, S. (2013). Implicit prosody and contextual bias in silent reading. *J. Eye Mov. Res.* 6, 1–14. Available online at: http://www.jemr.org/online/6/2/4

McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *J. Mem. Lang.* 38, 283–312. doi: 10.1006/jmla.1997.2543

Niebuhr, O. (2009). F0-based rhythm effects on the perception of local syllable prominence. *Phonetica* 66, 95–112. doi: 10.1159/000208933

Pickering, M., and van Gompel, R. (2006). "Syntactic parsing," in *Handbook of Psycholinguistics*, Vol. 2, ed M. Traxler (New York, NY: Academic Press), 455–503. doi: 10.1016/B978-012369374-7/50013-4

R Core Team (2014). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Sandalo, F., and Truckenbrodt, H. (2003). Some notes on phonological phrasing in brazilian portuguese. *DELTA* 19, 1–30. doi: 10.1590/S0102-44502003000100001

Savill, N., Lindell, A., Booth, A., West, G., and Thierry, G. (2011). Literate humans sound out words during silent reading. *Neuroreport* 22, 116–120. doi: 10.1097/WNR.0b013e328342ed27

Schmidt-Kassow, M., and Kotz, S. (2009). Attention and perceptual regularity in speech. *Neuroreport* 20, 1643–1647. doi: 10.1097/WNR.0b013e328333b0c6

Schwarzschild, R. (1999). Givenness, avoidF and other constraints on the placement of accent. *Nat. Lang. Semant.* 7, 141–177. doi: 10.1023/A:1008370902407

Snedeker, J., and Trueswell, J. C. (2004). The developing constraints on parsing decisions: the role of lexical-biases and referential scenes in child and adult sentence processing. *Cogn. Psychol.* 49, 238–299. doi: 10.1016/j.cogpsych.2004.03.001

Sorensen, T., Hohenstein, S., and Vasishth, S. (2015). Bayesian linear mixed models using Stan: a tutorial for psychologists, linguists, and cognitive scientists. arXiv: 1506.06201. Available online at: http://arxiv.org/abs/1506.06201

Spivey, M. J., and Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: modeling the effects of referential context and lexical frequency. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 1521. doi: 10.1037/0278-7393.24.6.1521

Stan Development Team (2014). *Stan Modeling Language User's Guide and Reference Manual [Computer Program], Version 2.2.* Available online at: http://mc-stan.org/

Stolterfoht, B., Friederici, A. D., Alter, K., and Steube, A. (2007). Processing focus structure and implicit prosody during reading: differential ERP effects. *Cognition* 104, 565–590. doi: 10.1016/j.cognition.2006.08.001

Sudhoff, S. (2008). "Focus particles in the German middlefield," in *The Discourse Potential of Underspecified Structures (Language, Context, and Cognition 8)*, ed A. Steube (Berlin/New York, De Gruyter), 439–459.

Swets, B., Desmet, T., Clifton, C. Jr., and Ferreira, F. (2008). Underspecification of syntactic ambiguities: evidence from self-paced reading. *Mem. Cogn.* 36, 201–216. doi: 10.3758/MC.36.1.201

Tantau, T., Wright, J., and Miletić, V. (2015). *The Beamer Class*. User guide for version 3.36.

Truckenbrodt, H. (2006). "Phrasal stress," in *The Encyclopedia of Languages and Linguistics*, Vol. 9, ed K. Brown (Amsterdam: Elsevier), 572–579. doi: 10.1016/B0-08-044854-2/04447-3

Vallduví, E., and Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics* 34, 459–520. doi: 10.1515/ling.1996.34.3.459

van Casteren, M., and Davis, M. (2006). Mix, a program for pseudorandomization. *Behav. Res. Methods* 38, 584–589. doi: 10.3758/BF03193889

van Gompel, R., Pickering, M., Pearson, J., and Liversedge, S. (2005). Evidence against competition during syntactic ambiguity resolution. *J. Mem. Lang.* 52, 284–307. doi: 10.1016/j.jml.2004.11.003

van Gompel, R., Pickering, M., and Traxler, M. (2001). Reanalysis in sentence processing: evidence against current constraint-based and two-stage models. *J. Mem. Lang.* 45, 225–258. doi: 10.1006/jmla.2001.2773

Vasishth, S., von der Malsburg, T., and Engelmann, F. (2013). What eye movements can tell us about sentence comprehension. *Wiley Interdiscipl. Rev. Cogn. Sci.* 4, 125–134. doi: 10.1002/wcs.1209

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

In the analyses reported in Section 2.3.3, we used the `lme4` package (Bates et al., 2014) and only fit varying intercepts models. This was because we could not fit maximal models due to convergence errors or boundary estimates of the correlation coefficients. In order to check that the effects hold up with a "maximal" model (Barr et al., 2013), we also fit Bayesian linear mixed models using the `stan_lmer` function in the `rstanarm` package (Gabry and Goodrich, 2016). It is almost always possible to fit a maximal model in the Bayesian setting because the weakly informative priors used in the modeling will ensure that the posterior distribution of a parameter will be centered around 0 if there is not enough data to estimate the true value of the parameter.

For the intercept, we used the prior $Normal(0, 6^2)$, and for the different comparisons, $Normal(0, 1)$. For the full variance-covariance matrices (subjects as well as items) we defined an LKJ prior (Stan Development Team, 2014) on the correlation matrix; see Sorensen et al. (2015) for more details. One way to interpret the Bayesian LMM is to examine the 95% uncertainty intervals and, in addition, to calculate the probability from the posterior distribution that the parameter is less than 0 ($P(b < 0)$). We will consider an effect to be present if the uncertainty interval doesn't contain 0. These effects are marked in bold in the tables below. The tables present modeling results for the three dependent variables (FPRT, RPD, TFT) in the four regions of interest (Subject, *auch*, Object, Verb), analogous to the conventional LMMs in Section 2.3.3.

log FPRT Subject

	Comparison	Mean	Lower	Upper	$P(b < 0)$
1	Intercept	5.5048	5.4543	5.554	0
2	Context	0.0073	−0.0191	0.0336	0.2968
3	RhythmL	−0.0261	−0.0572	0.0062	0.948
4	RhythmR	0.0154	−0.0127	0.043	0.1403
5	Context:RhythmL	0.0154	−0.0118	0.0432	0.132
6	Context:RhythmR	7e-04	−0.0265	0.0274	0.4902
7	RhythmL:RhythmR	−0.0105	−0.0374	0.0163	0.768
8	Context:RhythmL:RhythmR	0.008	−0.0212	0.0383	0.2938

log RPD Subject

	Comparison	Mean	Lower	Upper	$P(b < 0)$
1	Intercept	5.63	5.5783	5.6853	0
2	Context	0.0163	−0.0159	0.0486	0.1698
3	RhythmL	**−0.0345**	**−0.0687**	**−5e-04**	**0.9778**
4	RhythmR	0.016	−0.0206	0.0531	0.1948
5	Context:RhythmL	0.0013	−0.034	0.0373	0.4702
6	Context:RhythmR	5e-04	−0.0329	0.0339	0.484
7	RhythmL:RhythmR	−0.0051	−0.0389	0.0293	0.6182
8	Context:RhythmL:RhythmR	−0.0081	−0.0426	0.0266	0.6742

log TFT Subject

	Comparison	Mean	Lower	Upper	$P(b < 0)$
1	Intercept	5.7148	5.6607	5.7683	0
2	Context	2e-04	−0.0356	0.033	0.4985
3	RhythmL	**−0.0701**	**−0.1063**	**−0.0336**	**1**
4	RhythmR	0.013	−0.0225	0.0486	0.2372
5	Context:RhythmL	0.0059	−0.0297	0.0408	0.3582
6	Context:RhythmR	−0.0081	−0.0434	0.0277	0.6768
7	RhythmL:RhythmR	9e-04	−0.0348	0.0349	0.4768
8	Context:RhythmL:RhythmR	−0.0025	−0.0376	0.0326	0.5645

log FPRT *auch*

	Comparison	Mean	Lower	Upper	*P*(*b*<0)
1	Intercept	5.4225	5.3736	5.4707	0
2	Context	0.026	−0.0056	0.0577	0.0552
3	RhythmL	0.0085	−0.0179	0.0354	0.2668
4	RhythmR	−0.0126	−0.0398	0.015	0.823
5	Context:RhythmL	0.0251	−0.002	0.0516	0.035
6	Context:RhythmR	0.0037	−0.0224	0.0301	0.3988
7	RhythmL:RhythmR	−0.0059	−0.0343	0.0214	0.6585
8	Context:RhythmL:RhythmR	0.0158	−0.0105	0.0426	0.119

log RPD *auch*

	Comparison	Mean	Lower	Upper	*P*(*b*<0)
1	Intercept	5.5262	5.4727	5.5812	0
2	Context	0.0177	−0.0171	0.0535	0.1595
3	RhythmL	−0.003	−0.0358	0.0301	0.5672
4	RhythmR	0.0237	−0.0081	0.0566	0.072
5	Context:RhythmL	**0.0387**	**0.0074**	**0.0705**	**0.0082**
6	Context:RhythmR	0.0106	−0.0216	0.0447	0.2648
7	RhythmL:RhythmR	−0.017	−0.0505	0.0164	0.8445
8	Context:RhythmL:RhythmR	0.0217	−0.01	0.0532	0.088

log TFT *auch*

	Comparison	Mean	Lower	Upper	*P*(*b*<0)
1	Intercept	5.5691	5.5142	5.624	0
2	Context	0.0123	−0.0216	0.0464	0.2238
3	RhythmL	0.0031	−0.0284	0.0348	0.4278
4	RhythmR	−0.0146	−0.0494	0.0207	0.7925
5	Context:RhythmL	**0.0412**	**0.0073**	**0.0744**	**0.0088**
6	Context:RhythmR	0.011	−0.0214	0.0428	0.241
7	RhythmL:RhythmR	−0.0133	−0.0473	0.0202	0.7852
8	Context:RhythmL:RhythmR	0.0044	−0.028	0.0365	0.3888

log FPRT Object

	Comparison	Mean	Lower	Upper	*P*(*b*<0)
1	Intercept	5.519	5.463	5.5736	0
2	Context	−0.03	−0.0664	0.0081	0.9388
3	RhythmL	−0.0261	−0.0574	0.006	0.944
4	RhythmR	2e-04	−0.0352	0.035	0.489
5	Context:RhythmL	0.0259	−0.0056	0.0582	0.0548
6	Context:RhythmR	−0.0202	−0.0575	0.0168	0.8592
7	RhythmL:RhythmR	−0.0122	−0.0451	0.0207	0.7748
8	Context:RhythmL:RhythmR	0.0046	−0.027	0.0366	0.39

log RPD Object

	Comparison	Mean	Lower	Upper	*P*(*b*<0))
1	Intercept	5.7581	5.6855	5.8276	0
2	Context	−0.0255	−0.07	0.0205	0.8742
3	RhythmL	−0.0197	−0.0607	0.0228	0.8258
4	RhythmR	0.0026	−0.0385	0.0451	0.4522
5	Context:RhythmL	0.0173	−0.0265	0.0632	0.2195
6	Context:RhythmR	−0.0291	−0.0693	0.0121	0.9192
7	RhythmL:RhythmR	0.0025	−0.0411	0.0461	0.4552
8	Context:RhythmL:RhythmR	−0.0052	−0.0453	0.0344	0.6022

log TFT Object

	Comparison	Mean	Lower	Upper	*P*(*b*<0)
1	Intercept	5.7049	5.6316	5.7737	0
2	Context	**−0.0475**	**−0.0907**	**−0.0038**	**0.9802**
3	RhythmL	−0.0172	−0.05	0.0153	0.8408
4	RhythmR	−2e-04	−0.0396	0.0387	0.5062
5	Context:RhythmL	0.018	−0.0174	0.0538	0.1538
6	Context:RhythmR	−0.0151	−0.0556	0.0265	0.767
7	RhythmL:RhythmR	0.0042	−0.0367	0.0441	0.4232
8	Context:RhythmL:RhythmR	−0.0111	−0.048	0.0268	0.7185

log FPRT Verb

	Comparison	Mean	Lower	Upper	*P*(*b*<0)
1	Intercept	5.5593	5.4736	5.6382	0
2	Context	−0.0384	−0.0797	0.0047	0.9618
3	RhythmL	−0.0161	−0.0549	0.0236	0.7892
4	RhythmR	0.0292	−0.0184	0.0782	0.113
5	Context:RhythmL	−0.0095	−0.0477	0.0269	0.6828
6	Context:RhythmR	−0.0092	−0.0516	0.0325	0.6548
7	RhythmL:RhythmR	0.0039	−0.0447	0.053	0.436
8	Context:RhythmL:RhythmR	**0.0416**	**0.001**	**0.0814**	**0.0225**

log RPD Verb

	Comparison	Mean	Lower	Upper	*P*(*b*<0)
1	Intercept	6.1283	6.021	6.2321	0
2	Context	−0.0443	−0.0997	0.0121	0.94
3	RhythmL	0.0025	−0.0562	0.0632	0.4752
4	RhythmR	−0.0039	−0.06	0.0533	0.5495
5	Context:RhythmL	0.0181	−0.0345	0.0695	0.245
6	Context:RhythmR	0.0204	−0.0366	0.0766	0.2348
7	RhythmL:RhythmR	0.0107	−0.0448	0.0694	0.3618
8	Context:RhythmL:RhythmR	0.034	−0.0249	0.092	0.12

log TFT Verb

	Comparison	Mean	Lower	Upper	$P(b<0)$
1	Intercept	5.7045	5.612	5.8004	0
2	Context	−0.035	−0.0758	0.0065	0.9535
3	RhythmL	0.0134	−0.034	0.0588	0.281
4	RhythmR	0.0145	−0.0293	0.0577	0.2455
5	Context:RhythmL	0.007	−0.0308	0.0478	0.3675
6	Context:RhythmR	0.0061	−0.0357	0.0495	0.4012
7	RhythmL:RhythmR	0.0049	−0.0431	0.0544	0.4145
8	Context:RhythmL:RhythmR	0.0233	−0.0181	0.0659	0.1365

Rhythmic effects on syntactic encoding are limited to subsentential domains[1]

Gerrit Kentner & Isabelle Franz

Abstract

Does linguistic rhythm matter to syntax, and if so, what kinds of syntactic decisions are susceptible to rhythm? We investigated whether the propensity for rhythmic alternation between stressed and unstressed syllables affects the choice between introduced and unintroduced complement clauses in German. Apart from the presence or absence of the complementiser *dass* ('that'), these two sentence types differ with respect to the position of the tensed verb (verb-end/ verb second). We report on two recall-based sentence production experiments and two corpus studies - one on spoken and one on written language. Against our predictions, that were based on previously reported rhythmic effects on the use of the optional complementiser *that* in English, the experiments fail to produce compelling evidence for direct phonological influence on the structure of complement clauses in German. Only in the case of the written corpus did we find such an effect. In an attempt to integrate and interpret the sparse effects, we present an overview of studies which suggests rhythmic effects on syntactic encoding to be generally restricted to subsentential domains. We assume that, initially, sentence level syntactic projections (CP) are specified in the course of language production; their specification is in fact the prerequisite for phonological encoding to start. Therefore, prosodic effects may only touch upon the lower level categories that are to be integrated into the sentential projection, but not upon the syntactic makeup of the higher order projection itself.

[1] Acknowledgements:

1 Introduction

The sound of a sentence, its melody and rhythm is, in great measure, dependent on word choice, word order, and the choice of a particular syntactic construction. This paper is concerned with the question whether, and to what extent, the reverse holds as well: that is, whether the syntax of a sentence is dependent on prosodic aspects like melody and rhythm; or, put differently, whether speakers consider prosodic well-formedness when making syntactic decisions in language production.

Prosody's influence on syntax is most obviously attested in metered poetry; poets may tweak sentence structure to the benefit of sound and they may do so to an extent that would be unacceptable in normal speech, sometimes violating otherwise high-ranking syntactic rules (Fitzgerald, 2007; Kiparsky, 1975; Youmans, 1983). Similarly, given the importance of phonological form for persuasive speech (Menninghaus et al., 2015), speakers are known to adjust the syntax for the sake of prosody in rhetoric registers and they do so in both speech and writing (Anttila, 2017; Bolinger, 1957). Even in normal spontaneous language, prosodic influences on sentence structure have been reported. In what follows, we consider one specific prosodic feature, viz. the linguistic rhythm that is due to the distribution of stressed and unstressed syllables or of accented and unaccented words. The literature on this topic considers at least three differing ways in which linguistic rhythm may affect sentence structure (cf. Anttila, 2016, for a recent, more general, overview).

First, rhythmic alternation of stressed and unstressed syllables may be obtained by the inclusion or omission of optional elements. A language production experiment by Lee and Gibbons (2007) suggests that speakers use the unstressed optional complementiser *that* to maximise rhythmic alternation of weak and strong syllables, as it is more often produced when the top of the complement clause starts in a stressed (*Susan*) as opposed to unstressed (*Suzanne*) syllable (1). The unstressed complementizer followed by an unstressed initial of the proper name (1-b) would constitute a stress lapse (a sequence of two or more unstressed syllables), while a stress clash (a sequence of stressed syllables) ensues when the complementiser is omitted and a name with initial stress directly follows the stressed syllable of the embedding verb (1-c). Both stress lapse and stress clash deviate from alternation of stressed and unstressed and are thus considered comparatively dysrhythmic.

(1) a. Henry knew (that) {Súsan, Suzánne} washed the dishes.
b. Henry knéw **that Su**zánne washed the dishes. Stress Lapse
c. Henry **knéw Sú**san washed the dishes. Stress Clash

Secondly, in languages with sufficient flexibility concerning the word order, speakers have been shown to make use of this flexibility to ensure rhythmic alternation

(Schlüter, 2005; Rohdenburg, 2014). For example, Vogel et al. (2015) find that the syntactic placement of an inherently unaccented pronominal adverb (PA, underlined in (2)) is conditioned by the prosodic structure of the immediate environment, suggesting that speakers favour rhythmic (here: dactylic) linearisations like (2-a) and shun adverb placements which would lead to dysrhythmic word orders (2-b), i.e. those in which the alternation between stressed and unstressed syllables is less regular.

(2) a. Da wóllte der Péter Tomáten drin kóchen (rhythmic)
there wanted the Peter tomatoes PA-in cook
b. Da wóllte der Péter drin Tomáten kóchen (dysrhythmic)
there wanted the Peter PA-in tomatoes cook
Peter wanted to cook tomatoes in there

Thirdly, prosodic constraints may condition the choice between two or more (quasi-)synonymous sentence constructions. Anttila et al. (2010) and Anttila (2016) show that the choice between competing ditransitive constructions is dependent on the viability of the resulting prosody: specifically, double object constructions (8-a) are mostly illicit with verbs that have two stressed syllables (**She will dónàte Peter the book*). Moreover, speakers avoid ditransitive constructions involving a clustering of stressed syllables: Anttila et al. (2010) show that double object constructions involving a stressed goal (and concomitantly involving a stress clash, *gíve Jóhn the báll*) are clearly underrepresented compared to double object constructions involving an unstressed goal, as in *gíve him the báll.*

(3) Ditransitive constructions
a. Rita revealed him the truth (double object)
b. Rita revealed the truth to him (prepositional dative)
c. Rita revealed to him the truth (prepositional dative, direct object shifted)

Similar prosodic effects on the choice between genitive case and prepositional constructions for the expression of possessive or partitive attributes (e.g. *the mayor's house* vs *the house of the mayor*) have been reported by Shih et al. (2015).

The syntactic alternatives for both i) the English ditransitives (double object versus prepositional dative) and ii) the genitive attributes (genitive case vs prepositional genitive) involve different constructions that are semantically quasi-equivalent alternatives but that do not share a syntactic relationship. It is therefore not entirely clear whether the reportedly prosody-driven choice between these options affects syntactic operations *per se*, or whether it is better conceived as the selection between two precompiled syntactic frames, as advocated in construction grammar (Goldberg, 1995; Kay, 2002). In the latter case one might instead assume a prosodic effect on the selection from the 'lexicon of constructions', as it were, not so much a direct prosodic effect on syntactic computation or on syntactic relations.

Even in the cases where linguistic rhythm affects word order, it is not entirely clear whether this constitutes an effect on syntactic computation. Agbayani and Golston (2016) consider hyperbaton in Latin to be a type of word order alternation that involves phonological constituents rather than syntactic ones, and cannot be explained in purely syntactic terms. In this regard, the notion of 'prosodic movement' (Agbayani et al., 2011) embodies a conception of word order variation as phonological rather than syntactic, and this notion may well apply to cases like (2) in which the syntax is largely indifferent to the word ordering (see also Golston, 1995, for a similar point).

Finally, the prosody-driven inclusion or omission of optional elements need not be a prosodic effect on syntax, but merely a case of phonological ellipsis. This may be true of optional *that* in English complement clauses as there is no obvious syntactic difference between clauses with vs without *that* other than the presence or absence of the complementiser.

With these considerations in mind, the studies reviewed so far suggest that the observable rhythmic effects on sentence structure do not necessarily touch upon syntactic relations or syntactic computation per se. Instead, they may affect either the resulting phonological representation (the mere linearisation of syntactic constituents), or the choice among syntactically unrelated, possibly precompiled, constructions that happen to be quasi-synonymous.

However, there are cases of syntactic variability that appear to involve processes that may be more properly construed as syntactic in nature and that cannot be relegated either to the phonology or to the 'lexicon of constructions'. One example is the alternation between introduced (4-a) and unintroduced (4-b) complement clauses in German, which is the object of the study to be presented here. These structures are the German equivalent to English complement clauses with or without *that*.

(4) a. Die Studentin glaubt, dass Gisbert Techno hört.
The student thinks that Gisbert Techno listen-to.
b. Die Studentin glaubt, Gisbert hört Techno.
The student thinks Gisbert listens to Techno.

In contrast to its English analogue, this kind of alternation does constitute a difference in word order (verb-final subordinate clause versus verb-second structure), but the word order difference is, crucially, not reducible to phonology, as it affects syntactic constituents, not phonological ones. Also, the word order difference is not a simple one, as the structure in (4-a) requires the complementiser *dass* which is not licensed in verb-second structures like (4-b), i.e. there is a complementary distribution of subordinating conjunctions and verb-second order. On the other hand, one might say that (4-a) and (4-b) are different and independent constructions; but in contrast to the case discussed above (i.e. the English dative alternation), there is a systematic syntactic relationship between the two sentence structures. The conventional wisdom on these German structures holds that the verb-final order of a subordinate clause is the canonical or underlying word order. From this basic word order, other orders may be derived via verb movement and topicalisation (5) (see e.g. Thiersch, 1978; Wöllstein-Leisten et al., 1997, among many others).

(5) a. (dass) Gisbert Techno hört — underlying order
   (that) Gisbert Techno listen-to
 b. hört Gisbert Techno — verb movement → V-initial order
   listen-to Gisbert Techno
 c. Gisbert hört Techno — topicalisation → V2-order
   Gisbert listens to Techno

The alternation between complement clauses like (4-a) and (4-b) therefore constitutes a test case for the above-mentioned question whether linguistic rhythm affects syntactic processes in language production. Applying this question to the alternation concerning complement clauses in German is furthermore motivated by the above-mentioned language production experiment by Lee and Gibbons (2007) who found that the rhythmic environment affected the inclusion or omission of the optional complementizer *that* in English.

Section 2 provides more background on the syntax, semantics and pragmatics of introduced versus unintroduced complement clauses. In Section 3, we report on four experiments that largely fail to produce compelling evidence for a rhythmic effect on the choice between introduced and unintroduced complement clauses in German. We consider potential reasons for this in Section 4; specifically, we present a synoptic discussion of studies concerned with ostensible rhythmic effects on sentence structure and argue that rhythmic effects are restricted to subsentential domains of structure building. Section 5 concludes the paper.

2 Introduced vs unintroduced complement clauses

Finite complement clauses in German come in two varieties, viz. i) those that feature a complementizer (4-a) and ii) those that don't (4-b). The former are known as *introduced* complement clauses that display verb-final syntax, which is characteristic of most subordinate clauses in German. The latter are called *unintroduced* complement clauses for their lack of a subordinating conjunction (complementizer), or *dependent main clauses* (Auer, 1998) because their word order resembles the syntax of simple declarative clauses with the tensed verb in second position (V2). In the following, we focus on complement clauses that serve as sentential objects to a preceding verbal head and ignore sentence initial complement clauses or ones that are licensed by nouns or adjectives.

(4) a. Die Studentin glaubt, dass Gisbert Techno hört.
The student thinks that Gisbert Techno listen-to.
b. Die Studentin glaubt, Gisbert hört Techno.
The student thinks Gisbert listens to Techno.

The presence or absence of the complementizer usually does not affect the core meaning of the sentences; (4-a) and (4-b) are strictly synonymous. However, the literature on the subject notes several conditions for the choice between introduced vs unintroduced complement clauses.

For one thing, the syntax of finite complement clauses depends on the embedding verb. Several verbs license both introduced as well as unintroduced complement clauses but they do so to different degrees. While some embedding verbs equally appear with introduced and unintroduced sentential complements (e.g. *sagen, glauben*, 'say, believe'), other verbs hardly allow unintroduced complement clauses (this holds especially for factive verbs like e.g. *akzeptieren*, 'accept'). However, in general, the environments that license unintroduced complement clauses principally also license the variant with the complementizer.

The matrix clauses with embedding verbs have to be distinguished from so-called 'epistemic parentheticals' (Thompson and Mulac, 1991) or 'comment clauses' Brinton (2008). These are clauses that exclusively feature a restricted set of *verba putandi* in the first person singular (e.g. 'I think, I believe, I suppose'). Rather than serving as heads to the complement clause, these clauses are considered discourse markers or 'hedges' that signal uncertainty on the side of the speaker (Auer, 1998).

Apart from the lexical specifics of the embedding verb, the syntactic environment conditions the choice between introduced and unintroduced complement clauses. Negated matrix clauses decrease the likelihood of unintroduced complement clauses (Auer, 1998). The same holds when the embedding verb does not directly precede the complement clause. On the other hand, unintroduced complement clauses are more likely if they are set in conjunctive mood (Auer, 1998).

Auer (1998) claims that the relative pragmatic import of matrix clause and complement clause predicts the structure of the complement clause: If the complement clause contains presupposed information or information that the speaker considers known or discourse-given, it is more likely to be introduced with a complementiser; conversely, if the information in the complement clause is new or specifically relevant to the discourse, it is more likely to be realised as dependent main clause, i.e. without a complementiser. That is, Auer (1998) assumes that the syntactic subordination marker signals semantic or pragmatic subordination with respect to the main clause. This claim seems somewhat at odds with the findings by Ferreira and Dell (2000) and more recent research on English complement clauses: Ferreira and Dell (2000) show that speakers are more likely to omit the complementizer when the subsequent words were previously mentioned (i.e. known to the speaker). Jaeger (2010) found that the inclusion of the complementizer is more likely when the content of the complement clause is more informative, i.e. when its wording is less predictable from the context. Moreover, the frequency of the lexical material at the top of the complement clause is inversely correlated with the presence of the complementizer (Jaeger, 2010). Also, *that*-mention is strongly correlated with hesitations or disfluencies. A study by Hawkins (2004) suggests that the syntactic complexity or length of the complement clause increases the likelihood of the complementizer being present. Together, these findings suggest that, at least in English, the accessibility of the (lexical) material within the complement clause guides the production of this optional and syntactically redundant word; basically, speakers include *that* when they need to buy time to plan the subordinate clause, because its production turns out to be demanding. Other psycholinguistic studies find an effect of syntactic persistence. Speakers are more likely to produce the complementizer *that* when they did so in previous complement clause productions, but not necessarily when they produced *that* as demonstrative or relative pronoun, (Ferreira, 2003).

Weinert (2012) compares the use of introduced and unintroduced complement clauses in English and German; her data suggest that, in spoken conversation, more than 80 per cent of English complement clauses lack a complementizer, while the number is significantly lower in German (~ 60%). In both languages, the number of unintroduced complement clauses is lower in the written as compared to the spoken modality.

We are aware of two studies that specifically ascertain the role of phonology regarding the choice between introduced vs unintroduced complement clauses, viz. Jaeger (2006) and Lee and Gibbons (2007). Both studies are concerned with English optional *that* and both reach the conclusion that *that*-inclusion or omission is sensitive to phonological rhythm. The corpus study by Jaeger (2006) suggests that the inclusion of the generally unstressed *that* is significantly more likely in complement clauses the subject of which starts in a stressed as opposed to unstressed syllable. Jaeger's later, more comprehensive, follow-up study (Jaeger, 2010), however, does not include the rhythmic predictor any more. The likely reason for this is that, in this data set, the coding for stress at the top of the complement clause is confounded with the morpho-syntactic type of subject at the top of the complement clause (e.g. unstressed determiner or pronoun vs lexical noun with initial stress) and with the frequency of this word (e.g. high-frequency determiner/pronoun vs lower frequency lexical noun); these factors turn out to be strongly correlated with *that*-mention and most probably render the effects of stress and rhythm redundant.

The language production experiment by Lee and Gibbons (2007), on the other hand, controls for the morpho-syntactic type of the embedded subject and its frequency - it is always a disyllabic proper name with stress falling either on the first or on the second syllable.

Lee and Gibbons (2007), who found that the stress quality of the embedded subject affects *that*-mention, take their finding to support models of language production in which phonological encoding influences the grammatical encoding stage. This is especially noteworthy because psycholinguistic research has produced mixed, and rather little, evidence in favour of direct phonological feed-back to grammatical encoding (Vigliocco and Hartsuiker, 2002). For example, while Bock (1987) reports an effect of phonological priming on word order variation, neither Bock (1986) nor Cleland and Pickering (2003) find a comparable effect.

However, as mentioned above, for English, it remains doubtful whether unintroduced complement clauses really differ from those that present with the overt complementiser in syntactic terms. Alternatively, the complementiser *that* may be considered syntactically redundant and the presence or absence of it the business of the phonology. If the latter were true, with no tangible syntactic difference between the two types of complement clause, Lee and Gibbons' experiment would, after all, not constitute evidence for a phonological effect on grammatical encoding, but an effect that plays out exclusively within the phonological encoding stage.

In the following experiments we will examine whether the use of the German complementizer *dass* is systematically susceptible to the stress quality of the surrounding syllables. As discussed above, the choice between these variants involves a difference in word order that is not reducible to phonology and therefore clearly syntactic in nature. A positive result would strengthen the idea of bidirectional information flow between syntactic encoding and phonological encoding in language production. Anticipating the results, we did not find compelling evidence for such an interaction, against our predictions.

3 Experiments

3.1 Experiment 1

Experiment 1 is a conceptual replication of Lee & Gibbon's (2007) language production experiment, adapted to German, that closely follows their experimental protocol. Participants read sentence pairs silently in order to produce them afterwards in response to a recall cue. Sentence pairs consisted of a filler sentence as first stimulus and an experimental target sentence as second stimulus. Following Lee and Gibbons (2007), target sentences were simple matrix clause - complement clause structures that license the optional complementizer *dass* ('that') in German (6). The presence or absence of *dass* (together with the word order of the complement clause) and the stress quality of the surrounding syllables were systematically varied.

Assuming that participants vary their use of *dass* when producing the memorized sentences, we predicted that they do so in favor of an alternating rhythm. Thus, in sentences with an embedded trochaic subject (e.g. *Nadja*), the unstressed *dass* should be produced more frequently than in those with an iambic one (e.g. *Nadine*) as the latter would result in a stress lapse. Additionally, we predicted that an unstressed final syllable of the main verb would lead to more *dass* omissions, again in order to avoid a stress lapse. Conversely, the stressed monosyllabic embedding verb should result in more frequent use of the optional complementiser in order to avoid a stress clash.

3.1.1 Materials and Design

(6) a. Felix denkt, dass Nadine den Anzug gereinigt hat
b. Felix denkt, dass Nadja den Anzug gereinigt hat.
c. Felix denkt, Nadine hat den Anzug gereinigt.
d. Felix denkt, Nadja hat den Anzug gereinigt.
Felix thinks (that) {Nadja, Nadine} has cleaned the suit.

(7) a. Tim dachte, dass Nicole den Umschlag geöffnet hat.
b. Tim dachte, dass Nina den Umschlag geöffnet hat.
c. Tim dachte, Nicole hat den Umschlag geöffnet.
d. Tim dachte, Nina hat den Umschlag geöffnet.
Tim thought (that) {Nina, Nicole} has opened the envelope.

We constructed 32 items like (6) and (7), each involving a sentence frame with four conditions. Half of the presented sentences included the optional complementizer *dass* (6-a), (6-b), (7-a), (7-b), and concomitantly a verb final structure for the embedded sentence. The other half lacked the complementizer, featuring the tensed auxiliary *hat* in second position within the subordinate clause (6-c), (6-d), (7-c), (7-d). In each frame, the subject of the embedded complement clause was one of two disyllabic female first names, which differed in whether the first or the second syllable was stressed. As a consequence, in half of the items, the subject at the top of the subordinate clause started in an unstressed syllable (6-a), (6-c), (7-a), (7-c), and in the other half it started in a stressed one (6-b), (6-d), (7-b), (7-d). For the embedded subject we used 16 different name pairs for the 32 items so that every pair appeared twice in the set. Pairs were chosen that matched closely with respect to segmental content and frequency, as gleaned from the Leipzig Wortschatz corpus (http://wortschatz.uni-leipzig.de/). Further, in each frame the embedded sentence was constructed with one of 32 transitive verbs and their object in accusative case. The lexical verbs in the embedded clause were trisyllabic past participle forms (e.g. *gereinigt*, 'cleaned') and the auxiliary was always the monosyllabic *hat*, 'has'.

The main clause verbs were eight different verbs that license a sentential complement which may or may not be introduced by the comlementizer *dass* in spoken German. These eight main clause verbs were selected on the basis of Auer's (1998) study on complement clauses in spoken German. Auer (1998) identifies 13 high-frequency verbs that license both introduced and unintroduced complement clauses. For the purpose of our experiment, we included the ones that have both trochaic and monosyllabic word forms in their paradigms (*sagen, glauben, wissen, hoffen, hören, finden, meinen, denken*, 'say, believe, know, hope, hear, find, reckon, think').[2]

We systematically varied the stress position of the embedding verb as a between-items factor. Every main clause verb appeared four times in the set, two times as a monosyllabic (e.g. *glaubt*, 'believes') and two times as a trochee (e.g. *glaubte*, 'believed'). As a consequence, there were 16 items with monosyllabic (6) and 16 with disyllabic main verbs (7). Thirtytwo male proper names served as main clause subjects, half of them trochees and the other half monosyllabic ones, i.e. matrix clause subject and verb together consisted of exactly three syllables, as in the study by Lee and Gibbons (2007). That is, 16 of these matrix clauses ended in a stressed syllable (*Felix denkt*, 'Felix thinks') and 16 in an unstressed schwa-syllable (*Tim dachte*, 'Tim thought').

The experimental conditions thus ensured that, depending on the quality of surrounding syllables, the presence or absence of the unstressed *dass* yielded stress lapses or stress clashes at the clause boundary.

The 32 sets of materials were rotated around the experimental conditions in a latin square and mixed with fillers to yield four experimental lists. Each list contained 56 sentence pairs, 32 being experimental sentences paired with fillers and the other 24 pairs of fillers. The fillers were all unrelated to the experiment and did not contain the sentential complement structure. Each list contained exactly one version of each item, two appearances of each female first name and one appearance of each complement taking verb. Furthermore, each main verb appeared four times in each list, i.e. twice in present tense and twice in past tense. Accordingly, every male first name appeared one time in each list. Across the

[2]the verb *sehen* 'see' is also listed in Auer (1998); we excluded this verb because even the infinitive, the 1st person singular and the forms of the 1st and 3rd person plural are realized as monosyllables in most spoken registers: *sehe, sehen, sahen* [seː, seːn, saːn], so it is questionable whether this verb has trochaic forms at all.

lists, each version of each item occurred equally often and was paired with the same filler.

3.1.2 Participants

Thirty-two students from Goethe University Frankfurt, Germany, took part in the experiment. They identified as native speakers of German. In case of multilingualism, German was one of the mother tongues. Participants were paid 10 Euros for participation.

3.1.3 Procedure

Participants performed three practice trials followed by 53 trials. Each of the four lists was pseudorandomized eight times using Mix software (van Casteren and Davis, 2006) such that trials of the same condition or the same item had a minimum distance of two. This way, we created 32 different orders, one for each participant. Every participant started with the same 3 filler pairs as practice trials. Participants sat in front of a screen in a quiet room and were instructed to read and memorise two successively appearing sentences silently. Every sentence appeared for exactly five seconds. Then the recall cue for the first sentence appeared and participants had to produce the first sentence from memory. The cue stayed on the screen until participants indicated that they had finished by pushing a the enter key on a computer keyboard (60 seconds at most). After that, the second cue appeared immediately (consisting of three words, namely determiner-object-participle verb, e.g. *den Anzug gereinigt*, 'cleaned the suit') and participants hat to produce the second sentence from memory and push the key when they had finished. At this point, participants could take short breaks (again, 60 seconds at most) before they moved on to the next item. Participants were instructed to recall the names and words in the sentences as accurately as possible but were also encouraged to use their own words to express the meaning when they could not remember the exact wording. The presentation on the screen and the recordings of the participants' speech were programmed using psychopy software (Peirce, 2007). Additionally, the spoken productions were digitally recorded with an external recording device.

3.1.4 Scoring

All recordings were transcribed by an undergraduate student and scored by an additional student, adapting Lee & Gibbons' (2007) scoring scheme to the present design:

First, irrespective of presence/absence of the complementiser *dass*, a recall-based production was considered valid if it had one of the eight main clause verbs

	Final syll embedding verb		Initial syll embedded subject	
	stressed	unstressed	stressed	unstressed
introduced CC (with *dass*)	208	153	209	152
unintroduced CC (no *dass*)	57	60	78	39

Table 1: Distribution of introduced (upper row) vs unintroduced complement clauses (bottom row) in Experiment 1, broken down by stress quality of a) the final syllable of the embedding verb (left) and b) the initial syllable of the embedded subject (right).

immediately followed by a sentential complement with a first name as the embedded subject. Due to a very high amount of failed recalls (see results), we deviate from Lee & Gibbon's scoring scheme and did not exclude productions in which the word-prosodic structure of the embedding verb or of the embedded subject differed from the one in the stimulus sentence (Lee and Gibbons (2007) deemed those trials unusable). Rather, we scored for each valid production the prosodic status of the final syllable of the main verb (stressed or unstressed) and of the first syllable of the embedded subject (stressed or unstressed). Finally, we determined for each production whether it contained the complementiser *dass* or not.

3.1.5 Results and discussion

Out of the 1024 recordings, 478 trials (=47%) were considered valid according to the above criteria. This number is considerably lower than the number of the usable trials that Lee and Gibbons (2007) obtained in their experiment (678 out of 1024=67%). The reason for this discrepancy is unclear. We speculate that our filler sentences (the first sentence used in each pair of trials) are more complex than the ones used by Lee & Gibbons and correspondingly strained the memory to a higher extent. Table 1 lists the distribution of recalls with and without *dass* broken down by the stress quality of the surrounding syllables. As evident in Table 1, recalls involving *dass* are by far more common than those without. Also, there were considerably more valid recalls with trochaic names (n=287) as embedded subject compared to iambic names (n=191).

In order to estimate the influence of the rhythmic environment on the inclusion/omission of *dass* (and concomitantly on the syntactic structure of the subordinate clause), we fit a generalized linear mixed effects model (GLMM) with binomial link function, using the `lmer()` function (Bates et al., 2015) in the statistical computing environment R (R Core Team, 2014). The binomial dependent variable was the inclusion vs omission of *dass*. In line with the experimental design, we included as fixed effects i) the presence of *dass* in the stimulus sentence, ii) the stress status of the initial syllable of the proper name serving as embedded

Fixed effects:

	Estimate	Std. Error	z value	p value	
(Intercept)	1.7532	0.4382	4.001	6.31e-05	***
dassStimulus	0.9472	0.1625	5.828	5.60e-09	***
NameStress	-0.1451	0.1522	-0.954	0.340	
VerbStress	0.1497	0.1680	0.891	0.373	
NStress:VStress	0.1029	0.1492	0.690	0.490	

Table 2: Output of GLMM for Experiment 1.

subject (stressed vs unstressed), and iii) the stress status of the final syllable of the embedding verb (stressed vs unstressed) as between item variable; finally, iv) the interaction term of ii) and iii) was included as fixed effect. To avoid correlation of the fixed effects, we applied simple coding (orthogonal sum contrasts, with the two levels of each factor coded as -1 and 1, respectively). The experiment involves repeated measures for participant (n=32), item (n=32) and embedding verb (n=8); we therefore included these as grouping variables (or random effects). That way, the association between the data points coming from the same participant or item or embedding verb is taken into account in the estimation of the model's coefficients. Since the variance due to the items turned out to be minuscule, we report a more parsimonious model that only considers participant and embedding verb as random effects (this only marginally affects the coefficients of the fixed effects). The output of this model is shown in Table 2.

The positive estimate of the intercept of this model shows that recall-based productions involving *dass* were significantly more frequent. Furthermore, this model shows a significant effect of presence or absence of *dass* in the stimulus. That is, participants were significantly more likely to use *dass* when *dass* was present in the stimulus sentence. All other main effects and interactions do not significantly affect the inclusion/omission of *dass* in the recalled sentences. That is, there is no evidence in this experiment for a rhythmic effect on complementiser inclusion/omission, against our hypothesis that was derived from the results by Lee and Gibbons (2007). However, the validity of this experiment may be compromised because of the considerable data loss due to a great number of invalid recalls. We therefore set out to replicate this experiment with an easier distractor task in order to reduce memory load.

3.2 Experiment 2

Experiment 2 was a replication of Experiment 1 with a few alterations to reduce memory load. As in Experiment 1, participants read sentences silently in order to produce them afterwards in response to a three-word recall cue. Unlike Experiment

1, stimuli weren't sentence-pairs, but a combination of a target sentence as first stimulus and a simple arithmetic task.

3.2.1 Participants

Thirty-two students from Frankfurt and surrounding areas in Hesse, Germany, took part in the experiment. They identified as native speakers of German. In case of multilingualism, German was one of the mother tongues. Participants were paid 10 Euros for participation. None of them participated in Experiment 1.

3.2.2 Material and Design

We used the 32 items constructed for Experiment 1 as target sentences and also the 32 corresponding randomized lists to design this Experiment. In order to create an easier task, the second stimulus sentence of all 56 sentence pairs (i.e. the experimental sentences) of those lists were extracted and paired with a simple addition tasks while the order of the sentences was maintained. (Note that in Experiment 1 the first stimulus was always a filler sentence). As a consequence, stimuli were 56 combinations of a sentence as first stimulus and a mathematical task as second one. The arithmetic task was a simple addition involving as one addend the numbers 1, 2 or 3 and the second addend a double digit number (e.g. *2+34*) - the calculation never involved crossing a group of ten.

3.2.3 Procedure

The procedure was the same as in Experiment 1 with the exception that the second stimulus of each item was a simple arithmetic task. Participants were instructed to read and memorise one sentence silently and to solve the arithmetic task afterwards. Every sentence appeared for exactly five seconds. Then the additional task appeared and participants had to speak out loud the corresponding solution. The numbers stayed (60 seconds at most) on the screen until participants indicated that they had finished by pushing a button. Then the recall cue of the sentence appeared immediately and participants had to produce the sentence from memory and push the button when they had finished. At this point, participants could take short breaks (again, 60 seconds at most) before they moved on to the next item.

3.2.4 Scoring

The scoring scheme was the same as in Experiment 1.

	Final syll embedding verb		Initial syll embedded subject	
	stressed	unstressed	stressed	unstressed
introduced CC (with *dass*)	340	250	318	272
unintroduced CC (no *dass*)	214	169	202	181

Table 3: Distribution of introduced (upper row) vs unintroduced complement clauses (bottom row) in Experiment 2, broken down by stress quality of a) the final syllable of the embedding verb (left) and b) the initial syllable of the embedded subject (right).

Fixed effects:

	Estimate	Std. Error	z value	p value	
(Intercept)	1.21746	0.24915	4.887	1.03e-06	***
dassStimulus	2.65398	0.17447	15.212	< 2e-16	***
NameStress	0.03605	0.11063	0.326	0.745	
VerbStress	-0.01574	0.12481	-0.126	0.900	
NStress:VStress	-0.14407	0.10954	-1.315	0.188	

Table 4: Output of GLMM for Experiment 2.

3.2.5 Results and Discussion

Out of the 1024 recordings, 973 (=95%) were valid according to the above criteria. Table 3 depicts the distribution of recalls with vs without *dass*, broken down by the characteristics of the rhythmic environment (left panel: preceding syllable; right panel: following syllable). As in Experiment 1, productions involving introduced complement clauses (with the complementiser) outnumber those with unintroduced complement clauses. Also, participants were more likely to recall the embedded subject as trochaic rather than iambic, and the embedding verb as ending in a stressed as opposed to unstressed syllable.

In order to ascertain the effect of the rhythmic environment on complementiser We fit a GLMM with the same parameters as in Experiment 1. The output of the model is shown in Table 4.

As in Experiment 1, the significant intercept of the model shows a preference for productions involving *dass* over productions without. The significant main effect of *dass*-presence/absence in the stimulus clearly reveals that participants used *dass* more often when it was part of the stimulus sentence; all other main effects and the interaction term remain non-significant.

The lack of a significant effect of the rhythmic manipulation at the clause boundary in this experiment (and in the one reported above) does not hint at an influence of linguistic rhythm on *dass*-mention. To the contrary, the near-zero coefficients of the two rhythmic manipulations and of their interaction suggest that the choice between introduced and unintroduced complement clauses in German is largely unaffected by linguistic rhythm.

3.3 Experiment 3 - spoken language corpus

In order to validate the results of the language production experiments, we examined the *dgd* archive (`http://dgd.ids-mannheim.de`), the largest collection of spoken German corpora. For our purpose, we chose sub-corpora that contained transcriptions of unscripted speech by native speakers of Standard German only, namely the FOLK corpus and the Freiburg corpus, which comprise 2.6 million word tokens in 270 hours of speech in total. The dgd archive does not provide syntactic annotations other than part-of-speech (POS) tagging. Therefore, the search for relevant structures turned out to be rather time-consuming because the data had to be sifted, and the relevant prosodic features annotated, by hand.

3.3.1 Method

We searched for structures with an embedding verb directly followed by a complement clause with a proper name at the top. To find the relevant sentences, we looked up all bigrams with any of the previously identified, eight potentially embedding verbs (in any inflectional form) followed by a proper name, and the respective trigrams with the intervening complementiser *dass*. In all cases, we checked whether the proper name was indeed the subject of a subordinate clause headed by the embedding verb (especially for the searches without complementiser, this was very often not the case). Also, we discarded all instances in which the verb was part of a comment clause rather than a matrix clause. The search was hampered by the fact that the POS tag for proper names was identical to the POS tag for negations ("NE"), increasing the number of false hits substantially.

3.3.2 Results and discussion

In the end, the search yielded a rather small sample of 41 sentences. We coded the stress quality of the final syllable of the embedding verb and the initial syllable of the embedded subject proper name. Table 5 displays the instances of complement clauses with vs without *dass*, broken down by the stress quality of the embedding verb (left panel) and the embedded subject (right panel). All in all, the results suggests that unintroduced complement clauses are much more common than introduced ones in the corpus of spoken speech.

We employed Fisher's exact tests to test the (in)dependence of *dass*-mention and a) the quality of the preceding syllable (final syllable of embedding verb) and b) the quality of the following syllable (initial syllable of the subject). Neither

	Final syll embedding verb		Initial syll embedded subject	
	stressed	unstressed	stressed	unstressed
introduced CC (with *dass*)	7	1	8	0
unintroduced CC (no *dass*)	18	15	22	11
	Fisher's P=0.1201		Fisher's P=0.0831	
denk (2/6), find (0/3), glaub (3/17), hoer (0/1), hoff (0/1), mein (1/4), sag (0/0), wiss (2/1)				

Table 5: Distribution of introduced (upper row) vs unintroduced complement clauses (second row), broken down by stress quality of a) the final syllable of the embedding verb (left) and b) the initial syllable of the embedded subject (right). Results for Fisher's exact test are displayed below the 2×2 contingency tables. The list in the bottom part of the table displays the number of complement clauses found for each verb; the left number within the parenthesis stands for the number of introduced complement clauses, and the right number for unintroduced complement clauses.

test provides compelling reasons to discard the null-hypothesis that *dass*-use is statistically independent from the stress quality of surrounding syllables (cf. test statistics in Table 5).

Beyond the stress quality of the critical syllables at the clause boundary, we also coded the resulting rhythm as either alternating (rhythmic) or non-alternating (dysrhythmic). We consider as alternating rhythms those in which a) the final syllable of the embedding verb is stressed, followed by unstressed *dass* followed by initial stress on the embedded subject (e.g. *... denkt, dass Hans ...*, 'thinks that Hans'), or b) in which the final syllable of the verb is stressed, followed by an unstressed initial syllable of the embedded subject in unintroduced complement clauses (e.g. *... dachte, Hans ...*, 'thought Hans'); or c) in which the final syllable of the verb is stressed, followed by initial stress on the embedded subject (e.g. *... denkt, Maria ...*, 'thinks Maria'). Table 6 depicts the distribution of introduced and unintroduced complement clauses broken down by resulting rhythm (alternating or non-alternating). Again, Fisher's exact test does not disprove statistical independence of complementiser use and rhythm; therefore this distribution does not suggest that linguistic rhythm has an effect on the use of the complementiser in spontaneous, unscripted speech in German.

3.4 Experiment 4 - written language

The previous experiments do not hint at an effect of linguistic rhythm on the choice between introduced and unintroduced complement clauses in spoken German. However, we know that certain types of language use, styles or registers are especially prone to observe constraints on prosodic well-formedness. A host

	Resulting rhythm	
	alternating	non-alternating
introduced CC (with *dass*)	7	1
unintroduced CC (no *dass*)	20	13
	Fisher's P=0.2267	

Table 6: Distribution of introduced (upper row) vs unintroduced complement clauses (second row), broken down by the resulting rhythm (alternating vs non-alternating). Results for Fisher's exact test are displayed below the 2×2 contingency table.

of recent research suggests (see e.g. the collections of papers in Frazier and Gibson, 2015; Kentner and Steinhauer, 2017) that prosody plays a significant role in the processing of written language, in line with the *Implicit Prosody Hypothesis* (Fodor, 1998, 2002). This may seem paradoxical because written language lacks explicit cues to prosodic structure (e.g. Chafe, 1988). However, psycholinguistic research strongly suggests that, nonetheless, readers have immediate access to prosodic features like stress (Ashby and Clifton, 2005) when reading words, and they make use of this information when parsing sentences (Bader, 1996; Breen and Clifton, 2011, 2013; Kentner, 2012; Kentner and Vasishth, 2016). Writers are likewise known to structure their text in a way that aligns prominent words with sentence positions that are likely to receive a (nuclear) accent (Bolinger, 1957; Anttila, 2017). Furthermore, the process of writing itself appears to be accompanied by prosody. This is at least suggested by Fuchs and Krivokapić (2016) who show that pauses between key strokes during spontaneous writing are correlated with prosodic breaks in a read rendition of the same text.

Given the role of prosody in written language, we ask whether the choice between introduced and unintroduced complement clauses in the written modality remains unaffected by rhythmic-prosodic features of the words surrounding the clause boundary, as in the studies on spoken speech (Experiments 1, 2, and 3). This would suggest that this specific syntactic decision is a very stable one and cannot easily be undone in favour of phonological well-formedness – in contrast to other types of syntactic variation. However, an effect of linguistic rhythm in the written modality would not necessarily force us to commit to a bidirectional interaction between syntactic encoding and phonological encoding in language production because writers can revise their wording and the final text does not give away the spontaneous syntactic choice writers make, but only the end-result of a process that involves revisions. Still, an effect would at least show that writers do not tie themselves down to their initial syntactic decision concerning the form of the complement clause but consider the local rhythmic environment at the clause boundary when formulating complement clauses.

3.4.1 Method

To answer the above questions, we examine the TÜPP-D/Z corpus[3] that comprises all editions of the daily newspaper *die tageszeitung (taz)* from September 1st 1986 until 7 May 1999. This corpus contains 11.5 million sentences comprising 204.4 million word tokens. We searched for all tokens of the 8 embedding verbs that were immediately followed by a complement clause (with or without complementiser) with a proper name as clause-initial subject. This search yielded 2751 complement clauses, 1476 subordinate clauses with, and 1275 subordinate clauses without, the complementiser *dass*. Two student assistants hand-annotated the stress-status of i) the final syllable of the embedding verb (stressed or unstressed), and ii) the stress status of the initial syllable of the embedded subject proper name. While the verb-final syllable was either a stressed syllable or an unstressed schwa-syllable, the initial syllable of the proper name could bear primary stress (as in "Theodor" [ˈteːodɔɐ̯]), secondary stress (as in "Manuela" [ˌmanuˈeːla]) or remain unstressed (we assigned all syllables that directly precede the stressed syllable to this category, independently of the vowel quality, e.g. the first syllable in "Nicole" [niˈkɔl]). In a couple of instances the stress status could not be determined, either because the word shows variable stress ("Saddam" [ˈzadam] or [zaˈdam]) or because the name, and therefore its stress pattern, was unknown. The affected cases (65 or 2.4%) were discarded from further analysis, so 2686 cases remain, 1429 of which feature the complementiser.

As the use of the complementiser is assumed to be affected by specifics of the clause-initial subject (Jaeger, 2010; Roland et al., 2006), we determined the usage frequency and length of this word in order to consider these factors in the analysis. As an approximation for the phonological length, we simply took the number of (orthographic) characters. Furthermore, we devised a simple measure of frequency that is sensitive to the (rather narrow) temporal context of the corpus. We did this because we assume the frequency of names in a newspaper corpus to be heavily affected by the nature of the events reported - at least more so than in the case of generic words. To this end, we calculated the logarithm of the absolute frequency of the name within the corpus sample and used this as a predictor in our statistical model.

3.4.2 Results and discussion

Table 7 shows the distribution of complement clauses with vs without *dass*. This tabulation indicates that the rate of complementiser use is highest (.62) after

[3]URL: `http://www.sfs.uni-tuebingen.de/en/ascl/resources/corpora/tuepp-dz.html`

	Final syll embedding verb		Initial syll embedded subject		
	stressed	unstressed	main stress	2ndary stress	unstressed
introduced CC (with *dass*)	513	916	1058	117	254
unintroduced CC (no *dass*)	311	946	962	97	198
rate of introduced CCs	.62	.49	.52	.55	.56

Table 7: Distribution of introduced (upper row) vs unintroduced complement clauses (second row), broken down by the stress quality of a) the final syllable of the embedding verb (left) and b) the initial syllable of the embedded subject (right). The bottom row lists the respective rates of complementiser use.

embedding verbs ending in a stressed syllable, and lowest (.49) after verbs ending in a schwa-syllable. Moreover, the rate of complementiser use decreases as the degree of stress on the initial syllable of the embedded subject increases.

In the following, we report the statistical model with the best fit, as determined in a backward elimination process (Table 8). The "full" model included as fixed factor i) the stress quality of the final syllable of the embedding verb (VerbStress), ii) the stress quality of the initial syllable of the proper name serving as embedded subject (NameStress), iii) the interaction of VerbStress and NameStress, iv) the logarithmized frequency (FreqSubj) and v) logarithmized length of the embedded subject (LengthSubj). Note that, in contrast to the models above, NameStress was treated as a numeric variable rather than a categorical one because instead of a dichotomy between stressed and unstressed we are dealing with three degrees of stress (main stress, coded as "3"; secondary stress, coded as "2"; unstressed, coded as "1") in the sample at hand. The specific embedding verb (VerbLemma) was entered as random intercept into the model.[4]

From this "full" model, we discarded the interaction between VerbStress and NameStress, which turned out to be the non-significant coefficient closest to zero among the fixed factors, and, in a second step, the likewise non-significant factor LengthSubj. Neither of the two reductions deteriorated model fit, as verified using the `anova()` function in R.

The statistical model confirms predictions based on Jaeger (2010) and Roland et al. (2006): high-frequency embedded subjects promote *dass*-omission, i.e. unintroduced complement clauses. Furthermore, the model corroborates the suggestion that the stress quality of the syllables at both sides of the clause boundary affects the choice between introduced and unintroduced complement clauses. The

[4]We considered various models with different random effect structures. The most complex model was adjusted for differences in all the fixed factors. However, the complex models did not converge or suffered from nonsensical correlation parameters among the random effects (with values of, or near -1 or +1). We therefore opted for the simple model that calculates the intercept of the random effect only.

Fixed effects:

	Estimate	Std. Error	z value	p value	
(Intercept)	0.96497	0.44126	2.187	0.0288	*
NameStress	-0.11718	0.05853	-2.002	0.0453	*
VerbStress	0.43122	0.10046	4.293	1.77e-05	***
log(Freq)	-0.09250	0.03094	-2.989	0.0028	**

Table 8: Output of GLMM for Experiment 4.

significantly positive coefficient of VerbStress points in the predicted direction: unstressed *dass* is clearly preferred after a stressed syllable, confirming that writers strive for (implicit) rhythmic alternation at the clause boundary. Interestingly, the comma that necessarily marks the boundary between embedding verb and embedded clause according to German orthographic rules does not prevent the rhythmic effect from showing up; we suggests, in line with Truckenbrodt (2005), that this kind of clause boundary does not normally constitute a prosodic boundary in speech (although it certainly is a potential position for a prosodic break, see discussion in Truckenbrodt and Darcy, 2010). In fact, it is possible, and perhaps a means to promote reading fluency, for prosodic feet to straddle syntactic boundaries: In keeping with assumptions by Lahiri and Plank (2010), final stress on the embedding verb would then open a prosodic position for the unstressed *dass* to be integrated in, as in the case of (8-a). The same position would already be filled in the case of trochaic embedding verbs (8-b), hence *dass* would remain unparsed in the prosodic representation. Unparsed syllables, however, constitute a violation of regularities concerning prosodic structure, and renditions like (8-b) are therefore avoided.

(8) a. (Fe- lix) (denkt$_{<,>}$ dass) ...
(x -) (x -)
b. (Fe- lix) (dach- te$_{<,>}$) dass ...
(x -) (x -) -
Felix {a: thinks, b: thought} that ...

The coefficient of NameStress, on the other hand, points in the opposite direction: unstressed *dass* appears to be more likely the lower the degree of stress is on the initial syllable of the following name. This statistically significant, if weak, effect is clearly against predictions: We hypothesised, in line with Lee and Gibbons (2007), that initial stress on the embedded subject would promote *dass*-mention for reasons of rhythmic well-formedness, i.e. to maximise the alternation of stressed and unstressed syllables. Lest we speculate about the origins of this unpredicted effect, we note that, as in Experiments 1, 2, and 3, this experiment doesn't provide evidence for rhythmic optimisation of syntactic structure triggered by the stress quality of the embedded subject in German.

4 General Discussion

Before discussing in detail the potential reasons for the sparse outcome regarding the effect(s) of rhythm, and potential implications, we note two general results from the experiments that are largely in line with insights from the previous literature on the topic of complementiser use.

First, both Experiments 1 and 2 replicate familiar findings regarding structural priming (Ferreira, 2003; Ferreira and Dell, 2000; Lee and Gibbons, 2007): The participants showed a clear bias to mention the complementiser (and hence: a verb-final complement clause) when this conformed to the syntactic structure of the written stimulus.

Secondly, the results of all four experiments confirm a hypothesis suggested by previous findings by Ferreira and Dell (2000), Hawkins (2004), and Jaeger (2010) concerning English complement clauses, namely that complementiser use is inversely correlated with accessibility of the to-be-uttered material and with general ease of processing: Experiments 1 and 2 involved recall tasks and thus put the participants under cognitive pressure; in these experiments, productions involving the complementiser clearly outnumbered those with *dass*-less, unintroduced complement clauses. Also, the written corpus search in Experiment 4 yielded a higher rate of complement clauses including *dass*; written language is deemed to be less spontaneous and more effortful than normal speech, and this effort may be a factor (apart from the written norm) that promotes *dass*-mention. In contrast, the data from the spoken corpus (Experiment 3) shows the opposite: as in Auer (1998), unintroduced complement clauses were clearly more frequent in this spontaneously spoken data set, as it was likely produced with less effort when compared to the material from the other experiments.

4.1 Lack of rhythmic effect on *dass*-mention: potential reasons and implications

The two language production experiments and the two corpus studies presented here largely fail to support our initial predictions concerning the role of rhythm on the choice between introduced and unintroduced complement clauses in German. The only effect that points into the predicted direction is the effect of VerbStress found in Experiment 4: writers apparently favour the complementiser *dass*, and hence a verb-final complement clause, when the preceding verb ends in a stressed syllable. We interpret this effect to reflect the propensity for rhythmic alternation in written language. We hasten to add that this effect does not necessitate the possibility of direct interaction between grammatical encoding and phonological encoding in models of sentence production (as advocated in e.g. Vigliocco and Hartsuiker, 2002; Shih, 2014). Rather, this effect is explicable with recourse to the notion of a monitoring loop (Hartsuiker and Kolk, 2001; Levelt, 1983) that checks the output after the formulation stage of language production has been completed and may trigger stylistic repairs.

Apart from the effect of VerbStress in Experiment 4, the lack of a rhythmic effect in the direction predicted by Anttila et al. (2010), Shih (2014) or Vogel et al. (2015) in general, and Lee and Gibbons (2007) or Jaeger (2006) in particular is consistent across the four experiments. While it is problematic to argue on the basis of "absence of evidence", this consistency warrants commentary.

As alluded to in the introduction, the kind of syntactic variability studied here is qualitatively different from the kinds of syntactic variability studied by other authors. Specifically, the choice between a complement clause with versus without the overt complementiser in German necessarily involves the choice between verb-final (with complementiser (4-a)) versus verb-second structures (without complementiser (4-b)). No such word order difference is involved in the choice between English complement clauses with vs without complementiser.

The fact that, in English, the difference between the two constructions merely affects the presence or absence of *that* invites the assumption that this seemingly syntactic difference is, in essence, a phonological one: that is, both complement clause variants may provide a structural position for the complementiser which, in the case of *that*-less complement clauses, simply remains phonetically empty. If the presence or absence of *that* is indeed regulated by the phonological processing module, a rhythmic effect on complementiser use as reported by Lee and Gibbons (2007) is explicable and expectable without assuming an interaction between grammatical and phonological encoding.

In the following, we submit an admittedly ad hoc and speculative, but testable, approach that accounts for the difference in susceptibility to phonological influences between complement clause selection in German (putatively no phonological effect), and other kinds of syntactic variation that have been shown to be affected by phonological constraints. Initially, we illustrate the approach on the basis of the variation concerning introduced vs unintroduced complement clauses in German on the one hand, and the variation concerning English ditransitive constructions (double object versus prepositional dative) on the other. The distribution of the latter in large corpora of spoken English has been shown to be conditioned by phonological factors such as the requirement of rhythmic alternation (Anttila et al., 2010; Anttila, 2016).

(3) Ditransitive constructions

a. Rita revéaled him the trúth (double object)
b. Rita revéaled the trúth to him (prepositional dative)
c. Rita revéaled to him the trúth (prep. dative, direct object shifted)

The crucial difference between i) the choice among introduced vs unintroduced complement clauses, and ii) among ditransitive constructions, we suggest, lies in the different phrase structural levels involved.

The choice between introduced and unintroduced complement clauses in German affects the structure of the CP, i.e. a sentential category. The choice between the double object and the prepositional dative construction, in comparison, involves phrasal categories that are sub-sentential, and that concern the structure and relation of (nominal) arguments within the verb phrase (VP).

We argue that the encoding of sentential categories is less prone to be affected by rhythmic constraints than subsentential categories, on the basis of the following assumption: sentential categories need to be specified earlier in sentence production than subsentential categories; we assume that the syntactic encoding of sentential categories is in fact the prerequisite for phonological encoding to start, and that is why their syntactic encoding is largely immune to influences of stress and rhythm. While subsentential categories (e.g. nominal arguments to the verb) may be activated early in the course of sentence production, their syntactic integration requires at least the partial generation of a higher phrasal category as scaffold, as it were. They are therefore more likely to be affected by demands on the accruing phonological representation. This assumption is in keeping with a proposal by Ferreira (2000) who suggests that the formulation stage in language production involves, in a first step, the generation of basic syntactic frames or "elementary trees" (Frank, 1992) that consist of a simple predicate plus its extended projection, i.e. the positions for its arguments; only once the argument slots are being filled, e.g. with the corresponding NP treelets, the phonological encoding may begin.

If true, this explanation makes strong predictions about what kinds of syntactic variability are susceptible to phonological well-formedness conditions and what kinds are relatively immune. Testing this prediction is beyond the scope of this paper. However, we can at least, based on the available evidence concerning rhythmic-phonological effects on syntactic encoding, provide a plausibility check. To this end, we give a synopsis of relevant studies and categorise the available evidence as follows (see Table 9):

The first set of studies (upper section of Table 9) suggests a rhythmic effect on the realisation of optional elements.[5] As discussed above, a phonological effect on the mention of such elements is explicable without assuming bidirectional in-

[5] The study by Lohmann (2011) is unique in that the effect of linguistic rhythm on *to*-mention fails to reach conventional levels of statistical significance.

formation flow between phonological and syntactic encoding; instead, we assume that these elements are present in a syntactic representation even when they are phonologically empty. That is, the presence/absence of these elements is exclusively regulated in the phonological encoding stage.[6] Accordingly, the fact that the phrasal level affected is seemingly sentential in the case of Lee and Gibbons (2007), does not constitute counter-evidence to our proposal.

The second set (middle section in Table 9) consists of studies in which speakers/writers exploit word order flexibility in order to achieve a favourable rhythmic representation. The most prominent phenomenon in this set is the ordering of (mostly NP/DP) conjuncts (Benor and Levy, 2006; Lohmann, 2014; McDonald et al., 1993). The word order variations do not touch upon the semantics of these constructions; however, they may signal a higher degree of expressiveness (e.g. the determiner inversion in Schlüter, 2005) or possibly involve stylistic mannerisms (e.g. the verb cluster ordering reported in Vogel et al., 2015). The syntax appears to be indifferent to at least some of these word order alternations: For example, determiner inversion engenders split constituents (*quite a long report* ~ *[a [[quite long] report]]*) and thus violates rules concerning phrasal integrity.[7] Nevertheless, we assume that the phenomena in this set are relevant for the stage of grammatical encoding that is concerned with linearising the syntactic structure (the positional level according to Bock and Levelt, 2002). In this set, the phrasal levels affected by linguistic rhythm are invariably subsentential.

Finally, the third set (bottom section of Table 9) contains cases in which speakers/writers consider linguistic rhythm when choosing a particular syntactic construction. This set involves studies on genitive (Shih et al., 2015) and dative construction choice in English (Anttila et al., 2010) and the case of negated attributive adjectives (avoidance of stress clash: $^{*/?}$*a nót pópular person* but ✓ *a nót very pópular person*, Schlüter, 2005). These phenomena involve syntactic decisions that go beyond the mere ordering of constituents (possibly touching upon the functional level of sentence production, cf. Bock and Levelt, 2002). The phrasal levels affected by rhythm in this set are, again, subsentential.

All in all, given this synopsis, it seems plausible that effects of linguistic rhythm on syntactic encoding are restricted to subsentential levels of phrase structure. The

[6] The recent report by Zuraw (2015) on optional *de* in French presents a similar case: Zuraw shows that the selection of the (null-)allomorph ([də], [d], [∅]) is susceptible to the segmental environment. However the absence of [də] or [d] in the phonological representation does not imply a gap in the syntactic structure.

[7] Several studies (Agbayani and Golston, 2016; Agbayani et al., 2011; Golston, 1995; Rice and Svenonius, 1998; Zec and Inkelas, 1990) consider similar word order phenomena which do not seem to adhere to syntactic principles. Agbayani et al. (2011) coin the term "prosodic movement" because the displaced elements are prosodic rather than syntactic in nature. However, since these studies are not concerned with the effect of stress and linguistic rhythm, we refrain from discussing them here.

lack of a rhythmic effect in our experiments is consistent with this assumption. However, the nature, and the sparseness, of the data in Table 9 do not allow firm conclusions to be drawn. For one thing, there are, as far as we are aware, no other studies directly testing effects of linguistic rhythm on the production of higher syntactic levels. The present study on complement clause structure in German appears to be the only one. Moreover, several studies only consider written material. As mentioned above, written corpora are not very informative about the process of language production, they merely reflect the result of the process, and they do so in a modality that is often considered secondary to the spoken modality. Nevertheless, in the studies that consider written and spoken data, the findings are largely consistent. On a more general note concerning the current state of science communication, there is a very strong bias for positive results to be published, and negative results are hardly reported (Open Science Collaboration, , 2015, and references therein). Therefore, the list in Table 9 has to be taken with some caution.

Study	Phenomenon	Language	Phrasal Level	Data type
Schlüter (2015)	optional *to* infinitive	English	VP	written
Lohmann (2011)	optional *to* infinitive	English	VP	written
Wasow et al. (2015)	optional *to* infinitive	English	VP	spoken and written
Lee and Gibbons (2007)	optional complementiser	English	CP	recall experiment
Rohdenburg (2014)	preposition doubling	(Low) German	PP	written
Vogel et al. (2015)	position of pronomial adverb	German	VP	recall experiment
Vogel et al. (2015)	verb cluster	German	VP	recall experiment
McDonald et al. (1993)	NP-coordination	English	NP	recall experiment
Benor and Levy (2006)	NP/AP/VP-coordination	English	NP/AP/VP	written
Lohmann (2014)	NP-coordination	English	NP	spoken and written
Schlüter (2005)	determiner inversion	English	NP	written
Anttila et al. (2010)	dative construction	English	VP	spoken and written
Shih et al. (2015)	genitive construction	English	NP	spoken and written
Schlüter (2005)	negated attributive adjectives	English	NP	written

Table 9: Survey of studies examining rhythmic effects on sentence structure.

5 Conclusion

The experiments presented here were designed to ascertain the extent to which linguistic rhythm (i.e. the preference for rhythmic alternation of stressed and unstressed syllables) affects syntactic encoding in sentence production. To this end, guided by the example of Lee and Gibbons (2007), we tested whether the choice between introduced and unintroduced complement clauses in German is influenced by the immediate rhythmic environment at the clause boundary. Against the predictions that were derived from similar effects in the literature, we failed to find a rhythmic effect on sentence structure in two recall-based production experiments and in a study examining a corpus of spoken German. The inspection of a written corpus, however, revealed a preference for the inclusion of the unstressed complementiser *dass* – and hence: verb-final structure in the complement clause – when it was preceded by an embedding verb that ends in a stressed syllable; conversely, the omission of unstressed *dass* was favoured after unstressed syllables in the written corpus. Based on a synoptic view of relevant studies, we propose a taxonomy of different phenomena that have been claimed to show rhythmic effects on syntactic structure building. This taxonomy distinguishes rhythmic effects on a) the (non-)realisation of optional elements; b) word order; and c) on the choice of a particular syntactic construction. The (non-)realisation of optional elements can be interpreted as exclusively affecting the phonological encoding of a sentence, leaving the syntactic representation untouched. Word order and construction choice, however, do represent stages of syntactic encoding. Conspicuously, in the studies we revisited, the phrasal levels that were shown to be affected by rhythm, are always subsentential, i.e. concerning VP, PP, or NP-internal structure. We take this to suggest that the lack of a rhythmic effect in the case of German complement clause structures is due to their being of a higher syntactic order: they represent a sentential category, specifically the structure of the CP. We argue that only once the syntactic slots that the sentential categories keep available are being filled may the phonological encoding begin. Accordingly, we assume that decisions about the structure of sentential categories remain largely unaffected by rhythmic-phonological encoding effects in spontaneous language production.

References

Agbayani, B., Golston, C., 2016. Phonological constituents and their movement in Latin. Phonology 33 (1), 1–42.

Agbayani, B., Golston, C., Henderer, D., 2011. Prosodic movement. In: Washburn, M. B., McKinney-Bock, K., Varis, E., Sawyer, A., Tomaszewicz, B. (Eds.), Proceedings of the 28th West Coast Conference on Formal Linguistics. Cascadilla Press, pp. 231–239.

Anttila, A., 2016. Phonological effects on syntactic variation. Annual Review of Linguistics 2, 115–137.

Anttila, A., 2017. Sentence stress in presidential speeches. In: Talk given at the annual meeting of the German linguistic society (DGfS). Saarbrücken.

Anttila, A., Adams, M., Speriosu, M., 2010. The role of prosody in the English dative alternation. Language and Cognitive Processes 25 (7-9), 946–981.

Ashby, J., Clifton, C., 2005. The prosodic property of lexical stress affects eye movements during silent reading. Cognition 96 (3), B89–B100.

Auer, P., 1998. Zwischen Parataxe und Hypotaxe: abhängige Hauptsätze im gesprochenen und geschriebenen Deutsch. Zeitschrift für germanistische Linguistik 26 (3), 284–307.

Bader, M., 1996. Sprachverstehen: Syntax und Prosodie beim Lesen. Westdeutscher Verlag.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67 (1), 1–48.

Benor, S., Levy, R., 2006. The chicken or the egg? A probabilistic analysis of English binomials. Language 82 (2), 233–278.

Bock, J. K., 1986. Meaning, sound, and syntax: Lexical priming in sentence production. Journal of Experimental Psychology: Learning, Memory, and Cognition 12 (4), 575.

Bock, K., 1987. An effect of the accessibility of word forms on sentence structures. Journal of Memory and Language 26 (2), 119–137.

Bock, K., Levelt, W., 2002. Language production. In: Altmann, G. (Ed.), Psycholinguistics: Critical concepts in psychology. Vol. 5. Taylor & Francis, pp. 405–453.

Bolinger, D. L., 1957. Maneuvering for stress and intonation. College Composition and Communication 8 (4), 234–238.

Breen, M., Clifton, C., 2011. Stress matters: Effects of anticipated lexical stress on silent reading. Journal of Memory and Language 64 (2), 153–170.

Breen, M., Clifton, C., 2013. Stress matters revisited: A boundary change experiment. The Quarterly Journal of Experimental Psychology 66 (10), 1896–1909.

Brinton, L. J., 2008. The comment clause in English: Syntactic origins and pragmatic development. Cambridge University Press Cambridge.

Chafe, W., 1988. Punctuation and the prosody of written language. Written Communication 5 (4), 395–426.

Cleland, A. A., Pickering, M. J., 2003. The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. Journal of Memory and Language 49 (2), 214–230.

Ferreira, F., 2000. Syntax in language production: An approach using tree-adjoining grammars. In: Wheeldon, L. (Ed.), Aspects of language production. Psychology Press, pp. 291–330.

Ferreira, V. S., 2003. The persistence of optional complementizer production: Why saying that is not saying that at all. Journal of Memory and Language 48 (2), 379–398.

Ferreira, V. S., Dell, G. S., 2000. Effect of ambiguity and lexical availability on syntactic and lexical production. Cognitive psychology 40 (4), 296–340.

Fitzgerald, C. M., 2007. An optimality treatment of syntactic inversions in English verse. Language Sciences 29 (2), 203–217.

Fodor, J., 1998. Learning to parse? Journal of Psycholinguistic Research 27 (2), 285–319.

Fodor, J. D., 2002. Psycholinguistics cannot escape prosody. In: Proceedings of the 1st International Conference on Speech Prosody. Aix-en-Provence, pp. 83–88.

Frank, R. E., 1992. Syntactic locality and tree adjoining grammar: grammatical, acquisition and processing perspectives. Ph.D. thesis, University of Pennsylvania.

Frazier, L., Gibson, E., 2015. Explicit and Implicit Prosody in Sentence Processing: Studies in Honor of Janet Dean Fodor. Springer, New York City.

Fuchs, S., Krivokapić, J., 2016. Prosodic boundaries in writing: Evidence from a keystroke analysis. Frontiers in psychology 7.

Goldberg, A. E., 1995. Constructions: A construction grammar approach to argument structure. University of Chicago Press.

Golston, C., 1995. Syntax outranks phonology: evidence from Ancient Greek. Phonology 12 (3), 343–368.

Hartsuiker, R. J., Kolk, H. H., 2001. Error monitoring in speech production: A computational test of the perceptual loop theory. Cognitive Psychology 42 (2), 113–157.

Hawkins, J., 2004. Efficiency and complexity in grammars. Oxford University Press.

Jaeger, T. F., 2006. Phonological optimization and syntactic variation: The case of optional that. In: Annual Meeting of the Berkeley Linguistics Society. Vol. 32. pp. 175–187.

Jaeger, T. F., 2010. Redundancy and reduction: Speakers manage syntactic information density. Cognitive psychology 61 (1), 23–62.

Kay, P., 2002. An informal sketch of a formal architecture for construction grammar. Grammars 5 (1), 1–19.

Kentner, G., 2012. Linguistic rhythm guides parsing decisions in written sentence comprehension. Cognition 123 (1), 1–20.

Kentner, G., Steinhauer, K., 2017. Prosody in written language - acquisition, processing, emergence of norms. Frontiers Media SA, Lausanne.

Kentner, G., Vasishth, S., 2016. Prosodic focus marking in silent reading: effects of discourse context and rhythm. Frontiers in Psychology 7, 319.

Kiparsky, P., 1975. Stress, syntax, and meter. Language, 576–616.

Lahiri, A., Plank, F., 2010. Phonological phrasing in Germanic: the judgement of history, confirmed through experiment. Transactions of the Philological Society 108 (3), 370–398.

Lee, M.-W., Gibbons, J., 2007. Rhythmic alternation and the optional complementiser in English: New evidence of phonological influence on grammatical encoding. Cognition 105 (2), 446–456.

Levelt, W. J., 1983. Monitoring and self-repair in speech. Cognition 14 (1), 41–104.

Lohmann, A., 2011. Help vs help to: a multifactorial, mixed-effects account of infinitive marker omission. English Language & Linguistics 15 (3), 499–521.

Lohmann, A., 2014. English coordinate constructions. Cambridge University Press.

McDonald, J. L., Bock, K., Kelly, M. H., 1993. Word and world order: Semantic, phonological, and metrical determinants of serial position. Cognitive Psychology 25 (2), 188–230.

Menninghaus, W., Bohrn, I. C., Knoop, C. A., Kotz, S. A., Schlotz, W., Jacobs, A. M., 2015. Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. Cognition 143, 48–60.

Open Science Collaboration, , 2015. Estimating the reproducibility of psychological science. Science 349 (6251), aac4716.

Peirce, J. W., 2007. Psychopy – psychophysics software in python. Journal of Neuroscience Methods 162 (1), 8–13.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL `http://www.R-project.org/`

Rice, C., Svenonius, P., 1998. Prosodic V2 in Northern Norwegian. Manuscript, University of Tromsø.

Rohdenburg, G., 2014. Grammatical variation and the avoidance of stress clashes in Northern Low German. In: Cheshire, J., Stein, D. (Eds.), Taming the vernacular: From dialect to written standard language. Routledge, New York, pp. 93–109.

Roland, D., Elman, J. L., Ferreira, V. S., 2006. Why is that? Structural prediction and ambiguity resolution in a very large corpus of English sentences. Cognition 98 (3), 245–272.

Schlüter, J., 2005. Rhythmic grammar: the influence of rhythm on grammatical variation and change in English. Mouton de Gruyter, Berlin.

Schlüter, J., 2015. Rhythmic influence on grammar: Scope and limitations. In: Vogel, R., van de Vijver, R. (Eds.), Rhythm in Cognition and Grammar. A Germanic Perspective. de Gruyter, Berlin, pp. 179–206.

Shih, S., 2014. Towards optimal rhythm. Ph.D. thesis, Stanford University.

Shih, S., Grafmiller, J., Futrell, R., Bresnan, J., 2015. Rhythm's role in genitive construction choice in spoken English. In: Vogel, R., van de Vijver, R. (Eds.), Rhythm in Cognition and Grammar. A Germanic Perspective. de Gruyter, Berlin, pp. 207–233.

Thiersch, C., 1978. Topics in German Syntax. Ph.D. thesis, MIT, Cambridge, MA.

Thompson, S. A., Mulac, A., 1991. The discourse conditions for the use of the complementizer that in conversational English. Journal of Pragmatics 15 (3), 237–251.

Truckenbrodt, H., 2005. A short report on intonation phrase boundaries in German. Linguistische Berichte 203, 273–296.

Truckenbrodt, H., Darcy, I., 2010. Object clauses, movement, and phrasal stress. In: Erteshik-Shir, N., Rochman, L. (Eds.), The sound patterns of syntax. Oxford University Press New York, NY, pp. 189–216.

van Casteren, M., Davis, M., 2006. Mix, a program for pseudorandomization. Behavior Research Methods 38 (4), 584–589.

Vigliocco, G., Hartsuiker, R. J., 2002. The interplay of meaning, sound, and syntax in sentence production. Psychological bulletin 128 (3), 442.

Vogel, R., van de Vijver, R., Kotz, S., Kutscher, A., Wagner, P., 2015. Function words in rhythmic optimisation. In: Vogel, R., van de Vijver, R. (Eds.), Rhythm in Cognition and Grammar. A Germanic Perspective. de Gruyter, Berlin, pp. 255–276.

Wasow, T., Levy, R., Melnick, R., Zhu, H., Juzek, T., 2015. Processing, prosody, and optional to. In: Frazier, L., Gibson, E. (Eds.), Explicit and Implicit Prosody in Sentence Processing. Springer, pp. 133–158.

Weinert, R., 2012. Complement clauses in spoken German and English: Syntax, deixis and discourse-pragmatics. Folia Linguistica 43 (1), 233–265.

Wöllstein-Leisten, A., Heilmann, A., Stepan, P., Vikner, S., 1997. Deutsche Satzstruktur: Grundlage der syntaktischen Analyse. Stauffenburg Verlag, Tübingen.

Youmans, G., 1983. Generative tests for generative meter. Language, 67–92.

Zec, D., Inkelas, S., 1990. Prosodically constrained syntax. In: Inkelas, S., Zec, D. (Eds.), The phonology-syntax connection. University of Chicago Press, pp. 365–378.

Zuraw, K., 2015. Allomorphs of French de in coordination: A reproducible study. Linguistics Vanguard 1 (1), 57–68.

 Linguistics 2015; 53(5): 1233–1241

Reply

Gerrit Kentner*

Problems of prosodic parallelism: A reply to Wiese and Speyer (2015)

DOI 10.1515/ling-2015-0029

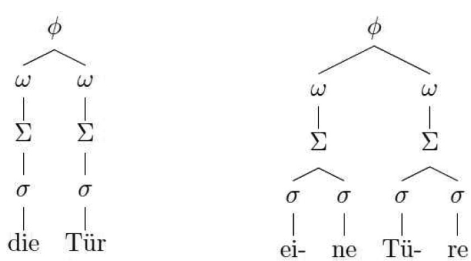
1 Prosodic parallelism and its virtues

In their recent contribution, Wiese and Speyer (2015) come forward with a very interesting proposal regarding the effect of supra-lexical prosody on word prosodic structure.

The proposal, the simplicity and elegance of which is captivating, is this: when given the choice, speakers strive for a rendition that maximizes prosodic parallelism; for two words that are prosodic phrase mates the foot structures are preferably parallel, i.e., the feet have the same number of syllables and stress pattern.

Wiese and Speyer build their account of prosodic parallelism on the analysis of schwa-zero alternations, examining a large corpus of written German. Specifically, they investigated several cases of nouns with apparently freely alternating monosyllabic and disyllabic variants like *Tür* – *Türe* 'door' (1) or *Tags* – *Tages* 'day$_{\text{GEN}}$' in the context of (preceding) monosyllabic or disyllabic determiners.

(1)



Using chi-square tests on bigram frequencies, they disprove statistical independence of the prosodic shapes of co-occurring determiner and noun, at least for

***Corresponding author: Gerrit Kentner,** Institut für Linguistik, Johann-Wolfgang Goethe-Universität, Norbert-Wollheim-Platz 1, 60629 Frankfurt am Main, Germany, E-mail: kentner@lingua.uni-frankfurt.de

the majority of the cases they investigated.[1] The results suggest that, more often than not, the number of syllables in the alternating noun corresponds to the number of syllables in the determiner, in line with the assumption of a constraint on prosodic parallelism.

This proposal is interesting and important for a number of reasons: here, I will raise three points that were not explicitly mentioned by Wiese and Speyer: First, the proposal connects well with (psycho)linguistic evidence to the effect that language users favor equal-sized constituents on many levels of phonological representation and processing (Fodor 1998; Ghini 1993; Myrberg 2013; Sandalo and Truckenbrodt 2003; Schweitzer et al. 2011; Selkirk 2000; Webman-Shafran and Fodor 2015, among others).

Second, there are phenomena that would defy proper analysis without recourse to a constraint on prosodic parallelism; these are cases in which the parallelism constraint appears to have a stronger influence than in the German schwa-zero alternations, in which parallelism is merely a tendency. Consider Standard Chinese, in which the productivity of N + N compounds and V + Obj combinations is strictly constrained by the number of syllables. As Duanmu (2012) shows, parallel prosodic structures with either two monosyllables (1 + 1) or two disyllables (2 + 2) are generally licit for both constructions. Crucially, however, for N + N compounds, non-parallel structures of the 1 + 2 type are mostly unacceptable. Similarly, for V + Obj phrases, the imbalanced pattern 2 + 1 is considered unacceptable (cf. Luo and Zhou 2010, for pertinent neurolinguistic evidence).

Another case demonstrating the influence of parallelism, again in German morphonology, is rhyme and ablaut reduplication (Kentner 2015a). These reduplications have a strict non-indentity requirement concerning base and reduplicant (*schickimicki*, **schickischicki* < *schick* 'posh'). Crucially, however, non-identity is confined to the segmental tier. That is, a difference between base and reduplicant concerning the prosodic shape is prohibited (**schischicki*, *??schickischick*), and it is this prohibition that strongly suggests the workings of prosodic parallelism.

1 Wiese and Speyer present a lot of confirmatory evidence but also discuss several cases in which the predictions of parallelism are not borne out. The effect appears to be modulated by frequency and lexicalization such that high frequency nouns and lexicalized determiner-noun combinations are less prone to be affected by parallelism. In the end, the variety of cases presented in favor of prosodic parallelism is convincing, but since a corpus analysis cannot possibly examine all potential cases in which parallelism is relevant, it remains unclear whether the sample is really representative.

Third, in indicating the force of parallelism in German morphonology, the proposal provides a striking argument for the hypothesis that poetic language – in which prosodic parallelism is prevalent – avails itself of constraints that are anyway active in normal language.

In the following, I will discuss three problems of Wiese and Speyer's application of the parallelism proposal to German schwa-zero-alternations. The first problem relates to the locus, or loci, of parallelism within the prosodic hierarchy. The second problem touches Wiese and Speyer's assumption that monosyllabic function words project a prosodic foot. The third, related, problem concerns the underlying assumption that prosodic phrases are built in accordance with the syntactic structure, to the effect that that determiner and noun are prosodic phrase mates. As I will show, all three problems make it difficult to evaluate Wiese and Speyer's proposal in relation to German schwa-zero alternations.

2 Problems of prosodic parallelism

2.1 The loci of parallelism within the prosodic hierarchy

Wiese and Speyer's approach to prosodic parallelism raises the question of which level of the prosodic hierarchy it may affect. This problem is best illustrated with cases that seemingly violate prosodic parallelism on one level of the hierarchy but at the same time ensure parallelism on another (higher) level. Consider (2) with the monosyllabic determiner *des* adjacent to the disyllabic form *Tages*. Under Wiese and Speyer's approach, the monosyllabic determiner should give rise to a preference for the monosyllabic noun *Tags* in this position. However, in the context of the phrase *am Ende*, the disyllabic variant leads to neatly parallel prosodic structures between the two phonological phrases (irrespective of whether the determiner is considered an unparsed syllable (2a) or promoted to a foot (2b) – on which more below). Conversely, the monosyllabic variant *Tags* would, in this context, undo the parallelism between the phonological phrases while achieving parallelism within one of them (but only under the analysis in (2b), which presupposes that the determiner does project a foot).

(2) a. *(am*$_{\sigma}$ *(Ende)*$_{\Sigma}$*)*$_{\Phi}$ *(des*$_{\sigma}$ *(Tages)*$_{\Sigma}$*)*$_{\Phi}$
b. *((am)*$_{\Sigma}$ *(Ende)*$_{\Sigma}$*)*$_{\Phi}$ *((des)*$_{\Sigma}$ *(Tages)*$_{\Sigma}$*)*$_{\Phi}$
at-the end the$_{\text{GEN}}$ day$_{\text{GEN}}$
'at the end of the day'

Wiese and Speyer's examination of the data is blind to effects of parallelism at higher levels of the prosodic hierarchy as their survey is focused on bigram frequencies, thus ignoring effects of the wider context. The general question evoked by cases like (2) is how parallelism at one level (between neighboring feet, say) interacts with parallelism at other levels (between phonological phrases). The case in (2) may be special because it is a set phrase used idiomatically – the higher-level prosodic parallelism (as between the phrases *am Ende* and *des Tages*) may possibly only develop in idiomatic contexts that guarantee a high co-occurence of the words involved. In any case, it appears that the predictive power of prosodic parallelism is undermined without further specification of the prosodic level(s) it exerts their influence on.

2.2 The prosodic status of function words

Wiese and Speyer's interpretation of the attested non-independence of prosodic shapes of determiner and noun presupposes that monosyllabic function words correspond to feet (cf. $(\text{des})_{\Sigma}$ $(\text{Tags})_{\Sigma}$, $(\text{ei.nes})_{\Sigma}$ $(\text{Ta.ges})_{\Sigma}$). However, the evidence for this assumption is weak at best.

While disyllabic determiners like *eine* 'a$_{\text{FEM}}$' or demonstrative pronouns like *diese* 'this$_{\text{FEM}}$' are clearly trochaic in citation form, it is unclear whether monosyllabic function words project a foot on their own. Indeed, in contrast to their English equivalents (*the* [ðə]), monosyllabic determiners in German have full vowels in citation form and fulfill the minimal requirement for phonological words in that they feature at least a long vowel (*die* [di:]) or a coda (*das* [das]) – that is two morae. Accordingly, assuming with the proponents of the prosodic hierarchy that wordhood implies the projection of at least a foot, these function words do correspond to feet. However, it does not seem outlandish to assume that the bimoraicity is not an inherent feature of the determiners but a consequence of their contextual isolation when uttered in careful pronunciation or in citation form – similar to the English determiner *the* that is phonologically enriched, as it were, when accented.

In actual speech, determiners and other function words are regularly subject to reduction (Hall 1999; Kabak and Schiering 2006; Vogel 2006; Wiese 1987). With respect to these parts of speech, reduction to moraless syllables may even be considered the norm rather than the special case. To be sure, this does not only hold for the monosyllabic forms but also for determiners that are disyllabic in citation form ([nəˈtyːɐ̯] <*eine Tür* 'a door') is orthographically represented as <'ne Tür> in more casual writing styles).

When considering the normal, reduced or even moraless pronunciation of the determiners in connected speech, prosodic parallelism cannot easily be made responsible for the correlation regarding the prosodic shapes of determiner and noun that was found in the written corpora. For moraless determiners cannot project a foot on their own, at least not under the standard assumption that a foot requires at least two morae. Instead, the determiner would have to be represented as an unparsed or cliticized syllable.

One may certainly argue that the reductions happen only by way of phonetic interpretation but do not affect the underlying phonological representation, in which even function words license feet (cf. the concept of Foot deletion or Defooting in Wiese [1987]). This argument is difficult to disprove, but it begs the question of whether feet are necessarily part of the lexical representation, which is especially doubtful in the case of function words. There is considerable evidence suggesting that feet are built in context, i.e., at least partially independent of the lexicon, and this will be reviewed in the following.

2.3 Prosodic phrasing and pedification beyond the word

Another problem concerning the workings of prosodic parallelism is rooted in the phonological phrasing that is assumed by Wiese and Speyer. As for the combinations of determiner and noun, Wiese and Speyer consider these words to be phonological phrase mates just as they form a syntactic unit together. As phonological phrase mates, the shape of the determiner may affect the prosodic shape of the alternating noun, according to Wiese and Speyer's parallelism proposal. The assumed phonological phrasing, however, doesn't go uncontested. Lahiri and Plank (2010) observe that (Dutch and English) speakers regularly choose a phrasing that maximizes the alternation of strong and weak syllables in a trochaic fashion. This, as Lahiri and Plank (2010) show, gives rise to prevalent mismatches between syntactic structure and prosodic phrasing. In German, like in Dutch or English, determiners (and other function words) tend to be weak monosyllables while nouns (and other lexical words) tend to be strong monosyllables or trochees (Eisenberg 1991). Following Lahiri and Plank, a phrase break is best placed before the strong noun with the weak determiner being adjoined to a preceding strong syllable (or remaining unparsed), thus establishing the preferred trochaic (and sometimes: dactylic) rhythm. The trochaic phrasing in (3) is one in which feet may straddle word boundaries. Correspondingly, the syntactic constituent Det-N (*'ne Flasche* in (3)) is broken up into two prosodic phrases.

(3) *(Hol mir) (mal 'ne) (Flasche) (Bier)*
bring me once a bottle beer
'get me a bottle of beer'

Following Lahiri and Plank (2010), determiners (and function words in general) are especially prone to reduction in this weak position that they are assigned to in phrasal prosody. Accordingly, determiner and noun are prosodified in a strictly asymmetric way, and, consequently, prosodic parallelism can hardly be made responsible for the facts regarding schwa-zero alternation on the noun, as reported in Wiese and Speyer. However, even under the trochaic phrasing assumed by Lahiri and Plank, there is a case for prosodic parallelism to be made. Crucially, under this proposal, parallelism may affect the prosodic shape of the function word rather than the shape of the noun. That is, determiners in prosodically weak positions are encliticized to the preceding strong syllable and thus reduced (cf. reduced determiner *'ne* in (3)), thereby creating a sequence of similar (parallel) trochees at the cost of syntax-prosody-isomorphism.

In Germanic, the preference for trochaic phrasing beyond the word is so strong that it evokes striking "slips of the ear", which result from the perceptual separation of strong syllables from preceding weak ones and the encliticization of weak syllables to preceding strong ones. Numerous studies by Cutler and colleagues (Cutler 2012) provide evidence for this rhythmic segmentation strategy in English and Dutch.[2]

Consider, in this regard, a German speaking child's reference to the toy character *Bob der Baumeister* 'Bob the builder', with *der Baumeister* functioning as the appositive attribute to the proper name: Whenever the 3-year-old is asked for the name of the character, the consistent answer, which defies proper alignment of syntax and prosody, is the optimal disyllabic trochee [ˈbɔp.də] - most likely a merger of the proper name and the following determiner.

To summarize, it appears that feet are built in context, and that function words tend to loose their status as foot licensors in connected speech (if they ever had this status). In other words, the strong propensity for trochaic rhythm regularly leads to (i) the reduction of determiners to weak syllables and (ii) the prosodic separation of determiner and noun.

3 Conclusion: Prosodic parallelism and the role of prosody in spoken and written language

The previous discussion casts doubt on two central premises of Wiese and Speyer's account, namely the presumed prosodic structure of function words

2 My own research on misperceived song lyrics (Kentner 2015b) strongly confirms the same tendency for German.

and the applied phonological phrasing. If, as I suppose, these premises do not hold, parallelism cannot explain the apparent prosodic correspondence of determiner and noun that Wiese and Speyer found in the written corpora. Moreover, Wiese and Speyer's study leaves open the question of which prosodic level(s) are potentially subject to parallelism and to what extent.

Anticipating (some of) the problems for their case, Wiese and Speyer explicitly state that written language (their object of study) does not represent actual speech but the "intended", i.e., unreduced pronunciation. It is certainly the case that the oral rendition of written language (i.e., read speech) usually contains far fewer reductions than spontaneous speech. Also, since, in writing, all orthographic words, irrespective of their syntactic role, are separated by blanks, one may assign each of them a similar (prosodic) status, with the blanks preventing cliticization, as it were. At the same time, reading research suggests that the implicit prosodic representation of the text in silent reading is essentially speech-like, even entailing fine phonetic detail (Ashby and Martin 2008; Chafe 1988; Filik and Barber 2011), but see de Ruiter (2015) who suggests that read speech and spontaneous speech have very different prosodic characteristics.

In any case, the specific role of prosody in written language and how it relates to speech prosody is only beginning to be explored (Breen 2014, for a review on the role of "implicit prosody" in reading). It is by no means clear how – in the absence of definite cues for phonological boundaries – written language is phrased phonologically. Worse still, given the un(der)specified prosodic structure of the written code, it is likely that readers apply a prosodic parse that differs from the one that the writer had in mind (provided that both readers and writers indeed apply prosodic phrasing to written language). Given this uncertainty, one cannot know which of the many conceivable prosodic representations is affected by prosodic parallelism and how.

Assuming that Wiese and Speyer's examination of schwa-zero alternations in written corpora reveals the workings of prosodic parallelism (as suggested by the authors), their results point to a fundamental difference concerning the role of prosody in written and spoken language. More concretely, the discrepancy between oral (speech) and written language leads to an apparently dialectical situation in relation to prosodic parallelism: On the one hand, effects of prosodic parallelism are seemingly observable in the written modality, which does not explicitly encode prosody. On the other hand, effects of this prosodic constraint remain largely undetectable because of regular reductions in the spoken modality, in spite of its rich prosodic code.

To conclude, even if prosodic parallelism is real in written language, we cannot know whether and how it affects spoken language.

References

Ashby, Jane & Andrea E. Martin. 2008. Prosodic phonological representations early in visual word recognition. *Journal of Experimental Psychology: Human perception and performance* 34. 224–236.

Breen, Mara. 2014. Empirical investigations of the role of implicit prosody in sentence processing. *Language and Linguistics Compass* 8. 37–50.

Chafe, W. 1988. Punctuation and the prosody of written language. *Written Communication* 5. 395–426.

Cutler, Anne. 2012. *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.

de Ruiter, Laura E. 2015. Information status marking in spontaneous vs. read speech in storytelling tasks: Evidence from intonation analysis using GToBI. *Journal of Phonetics* 48. 29–44.

Duanmu, San. 2012. Word-length preferences in Chinese: A corpus study. *Journal of East Asian Linguistics* 21. 89–114.

Eisenberg, Peter. 1991. Syllabische Struktur und Wortakzent. Prinzipien der Prosodik deutscher Wörter. *Zeitschrift für Sprachwissenschaft* 10. 37–64.

Filik, Ruth & Emma Barber. 2011. Inner speech during silent reading reflects the readers regional accent. *PloS one* 6. e25782.

Fodor, Janet Dean. 1998. Learning to parse? *Journal of Psycholinguistic Research* 27. 285–319.

Ghini, Mirco. 1993. Φ-formation in Italian: A new proposal. Toronto Working Papers in Linguistics. 12(2). 41–78.

Hall, T. Alan 1999. Phonotactics and the prosodic structure of German function words. In T. Alan Hall & Ursula Kleinhenz (eds.), *Studies on the phonological word*, 99–132. Amsterdam & Philadelphia: John Benjamins.

Kabak, Bariş & René Schiering. 2006. The phonology and morphology of function word contractions in German. *The Journal of Comparative Germanic Linguistics* 9. 53–99.

Kentner, Gerrit. 2015a. On the emergence of reduplication in German morphophonology. Frankfurt am Main: Ms. Goethe Universität.

Kentner, Gerrit. 2015b. Rhythmic segmentation in auditory illusions: The case of cross-linguistic mondegreens. Paper presented at the 18th International Congress of Phonetic Sciences, Glasgow, 10–14 August. Article ID 326.

Lahiri, Aditi & Frans Plank. 2010. Phonological phrasing in Germanic: The judgement of history, confirmed through experiment. *Transactions of the Philological Society* 108. 370–398.

Luo, Yingyi & Xiaolin Zhou. 2010. ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading. *NeuroImage* 49.2836–2849.

Myrberg, Sara. 2013. Sisterhood in prosodic branching. *Phonology* 30.73–124.

Sandalo, Filomena & Hubert Truckenbrodt. 2003. Some notes on phonological phrasing in Brazilian Portuguese. *Documentação de Estudos em Lingüística Teórica e Aplicada (DELTA)* 19(1). 1–30.

Schweitzer, Katrin, Kerstin Eckart, Petra Augurzky, Natalie Lewandowski & Grzegorz Dogil. 2011. A corpus-phonetic investigation of prosodic balance in German. Paper presented at the 17th Annual Conference on Architectures and Mechanisms of Language Processing (AMLaP 2011), Paris, 1–3 September. http://de.clarin.eu/de/dokumente/projektbibliographie.html (accessed 17 June 2015).

Selkirk, Elisabeth. 2000. The interaction of constraints on prosodic phrasing. In Merle Horne (ed.), *Prosody: Theory and experiment*, 231–262. Dordrecht: Kluwer.
Vogel, Ralf. 2006. Weak function word shift. *Linguistics* 44.1059–1093.
Webman-Shafran, Ronit & Janet Dean Fodor. 2015. Phrase length and prosody in on-line ambiguity resolution. *Journal of Psycholinguistic Research*. Advance online publication. DOI 10.1007/s10936-015-9358-2.
Wiese, Richard. 1987. Prosodic conditions on clitics. In Wolfgang U. Dressler, Hans C. Luschützki, Oskar E. Pfeiffer & John R. Rennison (eds.), *Phonologica 1984*, 331–338. Cambridge: Cambridge University Press.
Wiese, Richard & Augustin Speyer. 2015. Prosodic parallelism explaining morphophonological variation in German. *Linguistics* 53(3). 525–559.

New evidence for prosodic parallelism in German(ic) morphophonology

Gerrit Kentner[1]

[1]Goethe-Universität Frankfurt

kentner@lingua.uni-frankfurt.de

Abstract

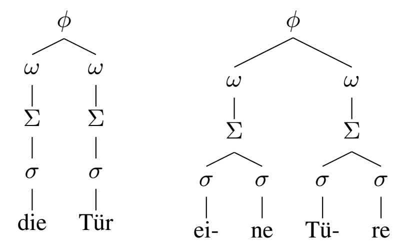
This paper presents two studies that make the case for prosodic parallelism as a factor in German(ic) word formation.

Index Terms: prosodic parallelism, prosody, morphology, prosodic morphology, Germanic

1. Introduction

In their recent contribution, Wiese & Speyer [1] (henceforth W&S) come forward with a very interesting proposal regarding the effect of supra-lexical prosody on word prosodic structure. The proposal, in nutshell, is this: when given the choice, speakers strive for a rendition that maximizes prosodic parallelism; for two words that are prosodic phrase mates, the foot structures are preferably parallel, i.e. the feet have the same number of syllables and stress pattern. W&S build their account of prosodic parallelism on the analysis of optional schwa, examining a large corpus of written German. Among other things, they investigated several cases of nouns with apparently freely alternating monosyllabic and disyllabic variants like *Tür* $\sim$ *Türe* ('door') or *Tags* $\sim$ *Tages* ('day$_{Gen}$') in the context of (preceding) monosyllabic or disyllabic determiners.

(1)



Using chi-square tests on bigram frequencies, they disprove statistical independence of the prosodic shapes of co-occuring determiner and noun. The results suggest that, more often than not, the number of syllables in the alternating noun corresponds to the number of syllables in the determiner as in (1), in line with the assumption of a constraint on prosodic parallelism.

In a response to W&S, I pointed out several problems concerning the case of determiner-noun sequences and the use of written corpora to ascertain the effect of prosodic parallelism [3]. Specifically, referring to common reduction phenomena in spoken speech, I questioned W&S's assumption that the determiner corresponds to a prosodic foot. A subsequent study [4] on the alternating adverbs *gern$\sim$gerne, selbst$\sim$selber, lang$\sim$lange* ('happily, oneself, for a long time') preceding various verb forms suggested that avoidance of stress lapse and stress clash, but not prosodic parallelism, account for the presence or absence of the schwa syllable on the adverb. Correspondingly, as it stands, the case for prosodic parallelism as a constraint on word or phrasal prosody appears to be weak.

Here, I present two case studies providing fresh evidence for the role of prosodic parallelism in German(ic) morphophonology. The cases suggest that a constraint on prosodic parallelism, albeit weak, is active on the word and phrasal level.

2. Parallel reduplication in German

The first case concerns rhyme and ablaut reduplications in German (2-a). These word formations are prime examples of prosodic morphology in that reduplication is only licit when native prosodic feet are involved [2]. Although rhyme and ablaut reduplication are mainly found in playful or facetious registers, they are subject to clear restrictions: Firstly, while reduplication is possible on the basis of monosyllables or trochees, reduplication with non-native feet or more complex foot structures are ungrammatical or at least clearly degraded (**Yvónnepivònne* $<$ *Yvónne, *Manuélapanuèla* $<$ *Manuéla*). Secondly, rhyme and ablaut reduplications observe a strict non-identity requirement regarding the segmental structure; base and reduplicant need to differ minimally, yielding the characteristic ablaut or rhyme. Crucially, however, non-identity on the prosodic level (2-b), (2-c) is illicit – the two feet involved in reduplication have to be strictly symmetric, i.e. parallel in shape: if the base is monosyllabic, the reduplicant must be monosyllabic. Conversely, when the base is disyllabic, the reduplicant has to be disyllabic, too.

(2)
- a. Mischmasch, Hickhack, Krimskrams, Schickimicki, Ilsebilse, doppelmoppel
  'mishmash, bickering, bric-a-brac, posh person, proper_name-RED, double-RED'
- b. $^{??}$Mischemasch, $^{??}$Hickehack, $^{??}$Krimsekrams, $^{??}$Schickimick, *Ilsebils, *doppelmopp
- c. *Mischmasche, *Hickhacke, *Krimskramse, *Schickmicki, *Ilsbilse, *doppmoppel

This requirement on reduplication is best captured with the constraint on prosodic parallelism. The data thus constitute evidence for its validity in German morphophonology.

3. Prosodic parallelism in coinages

For the second case study, (mostly English) coinages for musical genres from the website `everynoise.com` were examined. These coinages are names and as such a suitable test case. In contrast to generic words, names are not as open to morphological processes like inflection or derivation which would potentially alter the prosodic rendering.

Besides simplex words (e.g. *pixie*), these coinages are either phrases (e.g. *swedish metal*), or compounds/blends (e.g. *trip hop*). To ascertain the effect of prosodic parallelism, all dyadic coinages (n=714) listed in `everynoise.com` were scrutinised. While the majority of these was non-parallel in nature (e.g. *chicago house*), the subset involving only monosyllables and trochees as members of the dyad (n=498) did show a significant influence of prosodic parallelism (cf. Table 1) over and beyond a strong preference for monosyllabic constituents.

		right constituent	
		monosyll	trochaic
left constituent	monosyll	221	48
	trochaic	164	65

Table 1: *Cross-tabulation of coinages by prosodic shape of left and right constituent.*

A general linear model with binomial link function that was applied to this subset confirms that the prosodic shape of the left member of the dyad (usually the morphological or syntactic dependent) is not independent of the prosodic makeup of the morphological head in the right member ($z=2.611$, $p=0.009$). Moreover, the morphosyntactic status of the dyad (compound or phrase) significantly affected the prosodic shape of the left member ($z=5.364$, $p<0.001$) with a higher number of trochees in the case of phrases.

4. Conclusions

The two case studies suggest that, even though the effect of prosodic parallelism on optional schwa appears to be limited, it nevertheless systematically conditions the phonological makeup of complex words and phrases – at least as long as native prosodic feet (i.e. monosyllables or trochees) are involved.

5. Acknowledgements

Thanks are due to Marc Schwab who helped sieving and annotating the `everynoise.com` corpus.
Note that parts of the introduction are taken verbatim from [3].

6. Bibliography

[1] R. Wiese and A. Speyer, “Prosodic parallelism explaining morphophonological variation in German,” *Linguistics*, vol. 53, no. 3, pp. 525–559, 2015.

[2] G. Kentner, “On the emergence of reduplication in German morphophonology,” *Zeitschrift für Sprachwissenschaft*, vol. 36, no. 2, to appear.

[3] G. Kentner, “Problems of prosodic parallelism: A reply to Wiese and Speyer (2015),” *Linguistics*, vol. 53, no. 5, pp. 1233–1241, 2015.

[4] G. Kentner, *Schwa optionality and the prosodic shape of words and phrases*. Ms. Uni Frankfurt, 2016.

DE GRUYTER MOUTON
DOI 10.1515/tlr-2013-0009 — The Linguistic Review 2013; 30(2): 277–311

Gerrit Kentner and Caroline Féry

A new approach to prosodic grouping

Abstract[1]**:** This paper reports on two experiments concerning the prosodic realization and perception of various sentences with three or four coordinated names in German. The expression of prosodic boundaries, as evidenced by pitch and duration, is shown to signal the depth of syntactic embedding of the conjuncts and also the branching direction of the co-ordination structure. The results of the production experiment inspire a model of syntax-prosody mapping, which assumes that the strength of a prosodic boundary after a given constituent is a function of a) the syntactic relation to the following constituent and b) the depth of its syntactic embedding. Comparison reveals that the proposed model provides better predictions than other current approaches to prosodic boundary strength. The perception experiment indicates that listeners recognize recursively embedded coordination structures on the basis of the prosodic form of the sentence. We argue for a recursive representation of prosodic constituent structure at the level of the phonological phrase and above.

Keywords: prosodic phrasing, boundary strength, recursion, branching direction

Gerrit Kentner: Goethe-Universität Frankfurt, Institut für Linguistik
E-mail: kentner@lingua.uni-frankfurt.de
Caroline Féry: Goethe-Universität Frankfurt, Institut für Linguistik
E-mail: caroline.fery@gmail.com

1 Introduction

Coordinated names, like *Anna and Bill or Mary*, form a syntactically ambiguous structure, in the same way as an arithmetic procedure like 3 – 2 + 1, which can be resolved as 2 or as 0, depending on the order of the operations. In the case of coordinated names, the ambiguity concerns the branching direction and the level of syntactic embedding of the construction: either all three names may be on the

1 This work is part of the "Prosody in Parsing" project within the DFG's priority program 1234 "Phonetic and Phonological Competence". We thank Juliane Böhme, Caroline Magister, Daniel Quernheim, and Verena Thießen for their support in running the experiments and praat scripting. We gratefully acknowledge the discussions with, and comments by, Mara Breen, Anja Gollrad, Frank Kügler, Hubert Truckenbrodt, Michael Wagner, Duane Watson, and Shravan Vasishth. The paper has greatly benefitted from the suggestions of three anonymous reviewers.

same level of embedding (1-a), or two adjacent names may be grouped together to form a complex constituent that figures at the same level of syntactic embedding as the remaining simplex name ((1-b) and (1-c)). Depending on the kind of conjunction used, the different groupings may impinge on the truth value of a sentence the conjoined names are part of.

(1) a. [Anna or Bill or Mary]
 b. [[Anna and Bill] or Mary]
 c. [Anna and [Bill or Mary]]

Researchers have examined how different groupings of coordinated elements are realized prosodically (as for instance Ladd (1992) and Wagner (2005) for English, Schubö (2010) and Féry and Truckenbrodt (2005) for German). All authors have investigated phonetic differences in duration or pitch at conjunct boundaries and found a strong dependency between the prosodic realization and the syntactic place of the conjuncts in the coordination structure.

According to the results of previous research (e.g. Cooper and Paccia-Cooper, 1980; Lehiste, 1983; Gee and Grosjean, 1983), it may be considered verified that the prosodic boundary between adjacent constituents tends to be stronger the stronger the syntactic boundary between these constituents is. Correspondingly, prosodic boundaries are said to reflect syntactic structure. However, it is open to debate how close the match between syntactic and prosodic structure is.

We present results of a production and a perception experiment on various structures with coordinated names in German. It turns out that the expression of prosodic boundaries, as evidenced by pitch and duration, signals the depth of syntactic embedding of the constituents as well as the branching direction of the coordination structure. The results of the production experiment inspire a model of syntax-prosody mapping which assumes that the strength of a prosodic boundary after a given constituent is a function of a) the syntactic relation to the following constituent and b) the depth of its syntactic embedding. This model is compatible with accounts that allow a recursive representation of prosodic constituent structure at the level of the phonological phrase and above (Féry and Schubö, 2010; Ito and Mester, 2012; Ladd, 2008 [1996]; Wagner, 2005). A perception experiment with the same material indicates that listeners recognize embedded coordination structures on the basis of the prosodic form of the sentence, confirming that listeners are able to decode recursive syntactic structures on the basis of prosodic cues.

In Section 2, we review previous experimental and theoretical work on the prosodic expression of syntactic structure, and we introduce a new model which accounts for the prosodic expression of syntactic boundaries. The production experiment is reported on in Section 3. Based on the results of the production experiment, we evaluate our model and compare it with the predictive success of two existing models of prosodic boundary strength in Section 4. Section 5 presents the results of a perception experiment on the coordination structures. We conclude with a general discussion in Section 6, where we take up the issue of recursion in prosody as well.

2 Background and new proposal

2.1 Previous experimental work

There has been a keen interest in the psycholinguistic and phonetic literature as to how prosodic boundaries correlate with syntactic structure, especially in the case of structurally ambiguous sentences. Cooper and Paccia-Cooper (1980), Gee and Grosjean (1983), and Ferreira (1993) examine the placement as well as the strength of prosodic and intonational breaks in relationship to syntactic structure in speech production; Clifton et al. (2002, 2006) discuss the interpretation of prosodic boundaries with respect to sentence processing. See Watson and Gibson (2004) and Frazier et al. (2006) for summaries of previous research.

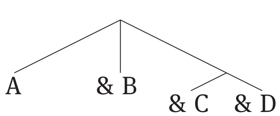
Speakers mark prosodic boundaries with characteristic acoustic cues: the duration of pre-boundary words is typically increased and there may be a period of phonetic silence; also, prosodic boundaries are characterized by deflections of pitch on the preceding syllable(s) (e.g. Cooper and Paccia-Cooper, 1980; Ferreira, 1993; Lehiste, 1983; Pierrehumbert, 1980; Price et al., 1991; Selkirk, 1984).

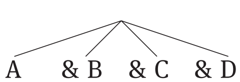
As for the relation between syntactic and prosodic boundaries, Watson and Gibson (2004) provide a model of prosodic boundaries called the Left hand side/Right hand side Boundary hypothesis (LRB), in (2), in which the sizes of the preceding and the following syntactic constituents are the predictors for the likelihood of intonational phrase boundaries. Intonational phrases are defined in Watson and Gibson (2004) as prosodic constituents of indeterminate length ending in a boundary tone and containing at least one syllable that receives a pitch accent (cf. Pierrehumbert and Hirschberg (1990)). Watson and Gibson's motivation for a model making reference to the size of constituents is related to processing demands: within a larger utterance, speakers need time to recover after particularly long constituents, and they need planning time for long upcoming constituents. The time needed for recovery and planning is provided by intonational phrase boundaries. Therefore, according to the LRB, the likelihood of an intonational break at any given word increases with the size of the surrounding constituents. The size of the left and right constituent are predicted to have an equal share in predicting the likelihood of an intonational boundary.

(2) The Left hand side/Right hand side Boundary Hypothesis (LRB, Watson and Gibson (2004))
The likelihood of an intonational boundary at a word boundary is a function of:
a. the size of the most recently completed constituent and
b. the size of the upcoming constituent if it is not an argument of the most recent head.

The LRB is shown to predict intonational phrase boundary location at least as well as, or even better than, more complex boundary strength models like Cooper and Paccia-Cooper (1980), Gee and Grosjean (1983) and Ferreira (1993). Watson and Gibson's own experiments, however, suggest that the LRB is too simplistic: their results show that the size of the preceding constituent has a much stronger influence on the likelihood of a boundary than the size of the upcoming one (see also Kentner (2007), who confirms this asymmetry for German).

Also, as Wagner (2005) observes, the LRB only predicts effects of adjacent constituents but cannot account for non-local effects of syntactic structure on boundary strength. In a production experiment, he found that simplex constituents such as A and B within a coordination structure like (3), which have a branching sister, are produced with longer duration than comparable simplex constituents that have only simplex sisters (4). Importantly, this also holds for simplex sisters that are non-adjacent to the complex constituent.[2]

(3)


(4)


2 Concurring with Wagner (2005), we consider coordinations of like categories (in this case: NPs and coordinations thereof), i.e. symmetric coordination. Correspondingly, n-ary branching trees are assumed to be appropriate syntactic representations when there are more than two conjuncts at the same level.

Accounting for such non-local effects, Wagner (2005) proposes an alternative model which relates the strength of prosodic boundaries to syntactic levels of embedding rather than the size of adjacent constituents. This is the Scopally Determined Boundary Rank (SBR) in (5).

(5) Scopally Determined Boundary Rank (SBR, Wagner (2005)):
If Boundary Rank at a given level of embedding is *n*, the rank of the boundaries between constituents of the next higher level is $n + 1$.

Although the predicted non-local increase in prosodic boundary strength due to embedding has been confirmed in Wagner's (2005) experiments, the SBR cannot easily account for the finding that the boundary strength also increases with the size or complexity of the surrounding constituents as predicted by the LRB and confirmed by the results of both Watson and Gibson (2004) and Wagner (2005). Moreover, as Wagner (2005) acknowledges, the SBR's success crucially depends on the use of different normalizing procedures depending on the various conditions tested.

Wagner's (2005) experiment on structures like (3) and (4) reveals another prosodic effect, which, however, neither the LRB nor the SBR succeed in predicting: the prosodic boundary after constituent C, if embedded as in (3), is significantly weakened relative to the boundary at the same position in the baseline pattern (4).

Given these problems of the LRB and SBR algorithms, we propose a new approach to the prediction of boundary strength based on two general principles that we call *Proximity* and *Similarity*.

2.2 The Proximity/Similarity model

We propose two general principles responsible for the interface between syntactic constituent structure and prosodic structure. These principles shape the expression of prosodic boundaries for the syntactic domain under consideration, i.e. a sentence or part thereof.

First, Proximity is inspired by a principle with the same name that Lerdahl and Jackendoff (1983) formulated in the context of musical grouping.[3] In Lerdahl and Jackendoff (1983), this principle is perception-oriented and amounts to the

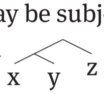
3 Lerdahl and Jackendoff's grouping principles are inspired by works in the tradition of Gestalt psychology (e.g. Wertheimer, 1938).

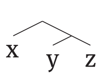
observation that two adjacent musical notes are perceived as belonging to different groups if the interval between them is large relative to other intervals in the vicinity. Here, Proximity operates on syntactic constituent structure, reflecting syntactic boundaries in prosodic structure. According to this principle, adjacent elements which are syntactically grouped together into one constituent should be realized in close proximity. Proximity between two elements is achieved by substantially weakening the prosodic boundary cues (segmental lengthening or boundary tone) on the first element. A corollary of Proximity is the opposite effect: adjacent elements not grouped together into one constituent should be realized with prosodic distance. As for Anti-Proximity, longer duration (final lengthening) and a higher boundary tone increase the distance to adjacent material to the right that is not part of the same immediate constituent. These effects are formalized in (6).

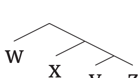
(6) Proximity
  a. The prosodic boundary at the terminal constituent x is weakened if the following terminal constituent y is the sister of x or dominated by the sister of x – unless x is immediately dominated by the root node of the domain under consideration.
  b. (Anti-Proximity): The prosodic boundary at the terminal constituent x is strengthened if the following terminal constituent y is not a sister of x.

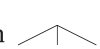
Note that (6) implies directionality because it is always the realization of the left of two elements that reflects whether the element to its right belongs to the same constituent or not. In other words, the prosodic expression of Proximity or Anti-Proximity on a lexical item only mirrors its syntactic relation to constituents to the right and not to those to the left.

There are four ways in which (6) may impinge on a lexical item:

(7) A lexical item x may be subject to
  a. Proximity (P) in 

  b. Anti-Proximity (A) in 

  c. both P and A in 

  d. neither P nor A (baseline) in 

In (7-a), x is subject to Proximity, since x and the following terminal y belong to the same immediate constituent to the exclusion of z. Anti-Proximity is shown in (7-b), where x does not belong to the same constituent as y and z. Proximity and Anti-Proximity have contradictory effects; a single lexical item may be subject to both when it is the left element of a larger embedded constituent, but the following terminal element is not its sister (7-c). In this case, we assume that the two effects cancel each other out.

The baseline representation in (7-d) corresponds to a list of lexical items with no hierarchical ordering. Here, all constituents are at the same level of embedding and are directly dominated by the root node. According to (6), the default prosodic break is neither strengthened nor is it weakened; instead, simple list intonation is predicted to apply.[4]

The second principle, Similarity, operates on the depth of syntactic embedding. It claims that constituents at the same level of embedding should be realized in a similar way, that is, they should be similar in pitch and duration, irrespective of their inherent complexity.

Similarity predicts prosodic adjustment of simplex elements as compared to complex constituents at the same level of embedding. More specifically, simplex elements are lengthened to approximate the duration of the complex constituent. This also holds for simplex elements that are non-adjacent to complex constituents if they are at the same level of syntactic embedding.

(8) Similarity
The prosodic boundary at the terminal constituent x is strengthened if a sister constituent of x is complex.

The two principles are predicted to interact to shape the prosody of syntactic structures.

While previous research has provided evidence for effects that may be explained in terms of Proximity and Similarity (e.g. Hunyadi, 2006; Wagner, 2005; Watson and Gibson, 2004), it is as yet unclear whether these principles overcome the aforementioned shortcomings of the LRB or SBR algorithms; if so, it is not obvious what the relative contribution of the two principles is, i.e. how much of the prosodic surface structure is attributable to the workings of Proximity and how much is due to Similarity. Moreover, a syntax-prosody mapping model that

4 We suggest that the characteristics of the default prosodic break depend on the structures under scrutiny. In the current case, the string of conjoined names makes up an intonational phrase that is separated by prosodic phrase boundaries after each name, where prosodic phrase is understood as a prosodic unit that contains one pitch accent.

makes use of Proximity and Similarity has to be clear about how these factors interact given that syntactic structures are subject to both.

To answer these questions, we conducted a production experiment designed to test the effects of (recursive) syntactic grouping on prosodic structure. Assuming that speakers do produce prosody that signals recursive syntactic embedding, it then remains to be verified whether listeners are able to deduce such nested syntactic structure from the prosodic form. This will be examined in a perception experiment.

In this paper, we aim at developing a model with Proximity and Similarity as main predictors. On the basis of the observed prosodic patterns we show that the performance of the Proximity/Similarity model is superior to that of the LRB, the SBR and a model combining both the LRB and SBR.

3 Production experiment

3.1 Method and material

The production experiment is based on Wagner's (2005) very similar experiment on the prosody of coordinate structures in English. The material consisted of different groupings of three or four conjoined proper names, all disyllabic and trochaic, like *Mila, Nino and Willi*. All groupings tested in the experiment are illustrated in (9) and (10), where N1 stands for the first name, N2 for the second name and so on. The conjunction *und* ('and') was always used inside of a bracket, and the conjunction *oder* ('or') outside of a bracket.[5] The structures 4.4 and 4.5 include embedded groupings, which are right-branching in the case of 4.4 and left-branching in the case of 4.5. As a result, we have three right-branching structures, 3.2, 4.2 and 4.4, and three left-branching structures, 3.3, 4.3 and 4.5.

(9)	3.1	N1 N2 N3	*Nino oder Willi oder Mila*
	3.2	(N1 N2) N3	*(Nino und Willi) oder Mila*
	3.3	N1 (N2 N3)	*Nino oder (Willi und Mila)*

5 We are aware that the use of different conjunctions may have had additional confounding effects (see Ladd (1992) and also Féry and Truckenbrodt (2005) for the effect of different sentence conjunctions in a sequence of three coordinated sentences). However, using only one type of conjunction would have led to very dull sentences. Given that the speakers were provided with explicit bracketing to mark the respective conditions, we think any nuisance effects stemming from the different conjunctions will be minor.

(10) 4.1	N1 or N2 or N3 or N4	*Nino oder Willi oder Mila oder Susi*
4.2	N1 or N2 or (N3 and N4)	*Nino oder Willi oder (Mila und Susi)*
4.3	(N1 and N2) or N3 or N4	*(Nino und Willi) oder Mila oder Susi*
4.4	N1 or (N2 or (N3 and N4))	*Nino oder (Willi oder (Mila und Susi))*
4.5	((N1 and N2) or N3) or N4	*((Nino und Willi) oder Mila) oder Susi*
4.6	(N1 and N2) or (N3 and N4)	*(Nino und Willi) oder (Mila und Susi)*

Participants were presented altogether 4 items from each of the nine conditions. The items were presented on screen one by one in randomized order. The grouping condition was made explicit by brackets and by a logical form. To trigger the target structure, a context plus a question was presented (a screen display is exemplified in (11)). Additionally, the context and question were presented auditorily over headphones once the screen display was shown.

(11) Context: Susi and Lena always go to the pool together, and Willi also does a lot of swimming.
Question: With whom do you want to go for a swim tomorrow?
Target: With (Susi and Lena) or Willi.
Logical Form: $(a \wedge b) \vee c$

The participants were 21 female students from the University of Potsdam, monolingual speakers of German in their twenties, coming from the Northern area of Germany. They were paid 6 Euros or got credit points for their participation. Recordings were made in an unechoic chamber on a DAT recorder. The participants were instructed to read the context carefully and to pay attention to the best way of realizing the groupings. They were given as much time as they wanted to utter the answer, and had the opportunity to correct themselves. If corrections were made, the last production of the item in question was taken. Altogether, 756 sentences were recorded and analyzed, 252 with three names (21 subjects × 3 conditions × 4 contexts), and 504 sentences with four names (21 subjects × 6 conditions × 4 contexts).

3.2 Measurements

An example of the realization is given in Figure 1.

The recordings were re-digitized from DAT at a sampling frequency of 44.1 kHz and 16 bit resolution. Every name as well as every conjunction were

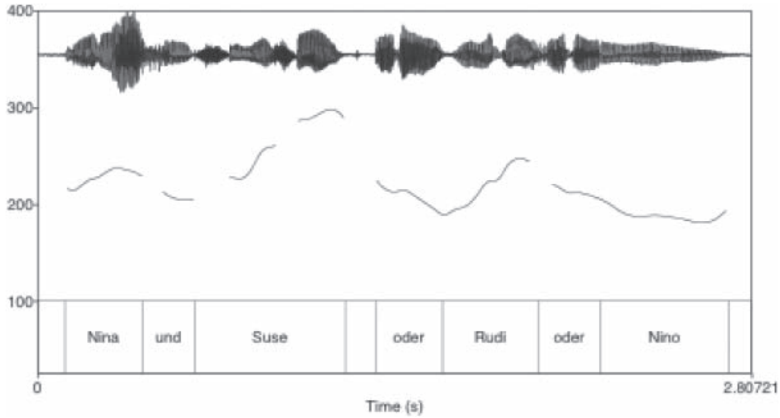


Fig. 1: Pitchtrack for condition 4.3.

labeled and delimited by a boundary set manually in an annotation tier in praat (Boersma and Weenink, 2009). We measured the duration of every name, of the pauses between names and of the conjunctions. As a measure of prosodic boundary strength, we summed the duration of each name and the following pauses, i.e. we considered the pauses part of the boundaries (see also Gee and Grosjean, 1983; Wagner, 2005; Wightman et al., 1992). A comparison with measurements without pauses did not reveal any relevant difference in the results. The analysis of pitch was conducted in praat, applying the smoothing algorithm (frequency band 10 Hz) to diminish microprosodic perturbations. Time-normalized contours were created by dividing up each constituent into five equal-sized intervals and by interpolating the aggregated mean F0 (in Hz) over speakers and sentences for each interval. All measurements were checked post hoc, and corrected manually when necessary (e.g. in the case of octave errors). Statistic analyses were performed using the statistical computing environment R.

3.3 Predictions

Based on earlier results from prosody research in German (Grabe, 1998; Féry and Kügler, 2008; Truckenbrodt, 2002, and others), some assumptions about the production of the expressions can be formulated. The realizations without grouping, 3.1 and 4.1, are taken as baselines and all other patterns are compared in relation to these baselines. In the baseline patterns without groupings, all names are expected to be of equal prominence and separated by boundaries of the same strength. Each name gets a pitch accent, which is expected to be rising (L*H) in non-final position and falling (H*L) in final position. L* and H* are the pitch accents, and the trailing tones H and L are the boundary tones of their respective domain. Pitch and duration of the final constituent are expected to be identical in all cases. In other words, we expect neutralization of the prosodic boundary at the end of all patterns, due to a final low boundary tone at the end of a declarative sentence. Another prediction is that, in the baseline, every high tone is downstepped relative to the preceding one, and no difference in duration occurs among the names.

If syntactic groupings are reflected in prosody, this is expected to happen by means of changed pitch accents, boundary tones and duration, the main intonational events. We derive our hypotheses about the prosodic realization of different syntactic groupings from the two general principles Proximity and Similarity.

As an example, the structures in 4.2 and 4.3 in (12) display one simple grouping of two elements into one constituent each.

(12) a. 4.2: *Nino oder Willi oder (Mila und Susi)*
b. 4.3: *(Willi und Mila) oder Susi oder Nino*

There are three constituents at the top level in these conditions, two simplex ones and a complex one. The simplex names are predicted to be lengthened and thus adjust to the duration of the complex constituent in order to achieve similarity across constituents at the top level. In addition, as predicted by Anti-Proximity, the element outside of but left-adjacent to a grouping should exhibit a stronger prosodic boundary (cf. *Willi* in (12-a)). The same applies to the rightmost name of a grouping (cf. *Mila* in (12-b)). The left elements of groupings are expected to show weaker prosodic boundary cues in order to fulfill Proximity (*Mila* in (12-a), *Willi* in (12-b)). To sum up, Proximity and Anti-Proximity should have local effects: weakening of the left and strengthening of the right element within a grouping, as well as strengthening of simplex elements that are left-adjacent to a grouping. Similarity implies that syntactic grouping has non-local effects as well: compared to the baseline, all simplex elements that have a complex sister should be lengthened (even those that are not adjacent to groupings). The different effects of Proximity (P), Anti-Proximity (A) and Similarity (S) are tabulated for each condition and each non-final name in Table 1 for the conditions with three names, and in Table 2 for the conditions with four names.

Table 1: Non-final names subject to Proximity (P), Anti-Proximity (A) and Similarity (S) in conditions with three names

	N1	N2
3.1 N1 or N2 or N3	–	–
3.2 (N1 and N2) or N3	P	A
3.3 N1 or (N2 and N3)	A,S	P

Table 2: Non-final names subject to Proximity (P), Anti-Proximity (A) and Similarity (S) in conditions with four names

	N1	N2	N3
4.1 N1 or N2 or N3 or N4	–	–	–
4.2 N1 or N2 or (N3 and N4)	S	A,S	P
4.3 (N1 and N2) or N3 or N4	P	A	S
4.4 N1 or (N2 or (N3 and N4))	A,S	P,A,S	P
4.5 (N1 and N2) or N3) or N4	P	A	A,S
4.6 (N1 and N2) or (N3 and N4)	P	A	P

3.4 Results for three names

The results for duration and pitch are shown simultaneously in Figure 2. In the description of the pitch contours, we concentrate on the high tones on the names themselves, and largely ignore the conjunctions, which behave as transitions between the names. The low tones are also discarded in the discussion. The base-

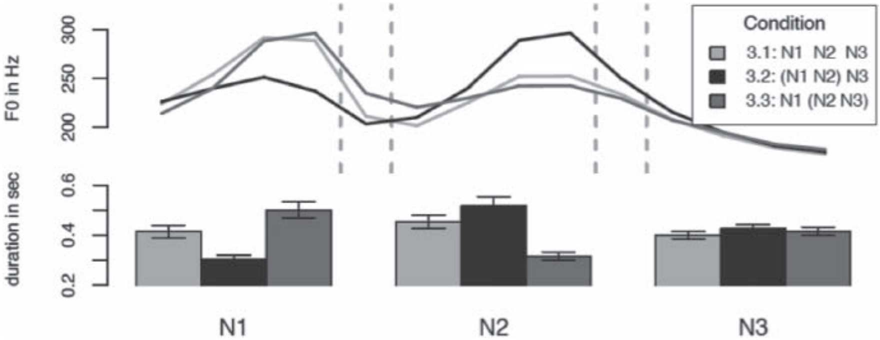


Fig. 2: mean pitch in Hz and mean duration in ms of the conditions with three names.

line pattern 3.1 (light grey) presents downstep between N1, N2, and N3. However, N3, the final name, is neutralized in all patterns, and will not be considered any further. Pattern 3.2 (black) shows an important difference compared to the baseline: N1's high tone is clearly lowered when compared to the baseline, while N2 has a higher pitch value (upstep), reaching a level comparable to N1 of the baseline condition. By contrast, the tonal pattern of 3.3 (dark grey), a right-branching structure, is very similar to that of the baseline 3.1. They both have a high N1 and subsequent downstep on the further two names. N1 in 3.3. is not significantly higher than N1 in the baseline condition. However, the N2 of pattern 3.3 is slightly lowered as compared to the baseline condition 3.1. As a result the difference in pitch (i.e. the amount of downstep) between N1 and N2 is larger in 3.3 than in 3.1. Comparing the high tones across conditions, a mirror-image relation between the left-branching condition 3.2 and the other conditions is apparent: the upstepped H-tone of N2 in 3.2 approximates the height of N1 in the other conditions. Conversely, the height of N1 in condition 3.2 closely resembles the height of the downstepped H-tones on N2 in the other conditions.

As for duration, the three names of the baseline pattern 3.1 (light grey columns) display small differences; the slightly longer duration of N2 (mean difference compared to N1 is about 40 ms) is significantly different from N1 ($t = 3.8$, $p < 0.001$). We return to this effect in the discussion (see Section 3.6). Compared to the baseline, pattern 3.2 (black) has a significantly shorter N1 (a group-initial element) and a significantly longer N2 (a group-final element). In contrast, in pattern 3.3 (dark grey), N1 (simplex element, left-adjacent to a grouping) is longer while N2 (group-initial element) is shorter than the baseline. We also see that N3's duration is neutralized. Indeed, this neutralization of the last name is persistent in all conditions, as we will see, both in duration and in pitch.

To sum up the three-name conditions, pitch and duration deliver equivalent results in that higher pitch on non-final names generally coincides with longer duration and lower pitch patterns with shorter duration. The pitch tracks reveal an interesting asymmetry: The right-branching pattern (3.3) has a striking resemblance to the baseline – both have a downstep pattern. But the left-branching pattern (3.2) has a different shape, namely a lower pitch on N1 and a clear upstep on N2. Both patterns with groupings clearly differ from the baseline with respect to duration. The first element of a grouping is always shorter than in the baseline, and the last element of a grouping is always longer than in the baseline (except in N3 because of final neutralization). These results are in line with the general principles of Proximity, Anti-Proximity and Similarity: names that are affected by Anti-Proximity and Similarity express a stronger prosodic boundary while the ones that are subject to Proximity are clearly shortened and lowered in pitch compared to the baseline.

3.5 Results for four names

In this section, we compare the realizations of the baseline 4.1 to the various conditions with groupings 4.2 to 4.6. An overview of all results on pitch and duration is given in the plots depicting difference scores between the baseline and other conditions with 95% confidence intervals in Figures 8 and 9 below.

First, the Figures 3 and 4 show the results for the right-branching conditions 4.2 and 4.4 as compared to the baseline condition 4.1. As was the case for the three-name patterns, the discussion for pitch concentrates on the relationship between the high tones of names. In the right-branching structures, 4.2 and 4.4, and in the baseline 4.1, there is downstep throughout. The general impression is that 4.2 and 4.4 have roughly the same shape as the baseline. However, in 4.2 and 4.4, N3 is somewhat lower than in the baseline. Correspondingly, the downstep between N2 and N3 is also larger than in the baseline, due to the fact that N3 is the

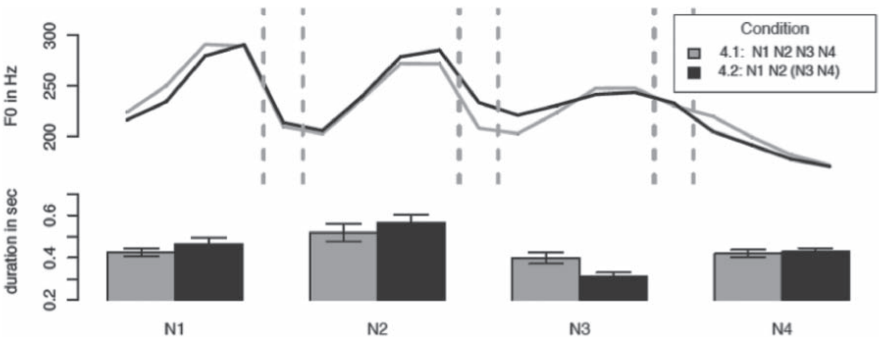


Fig. 3: Comparison of simple right-branching condition (black) with baseline (grey).

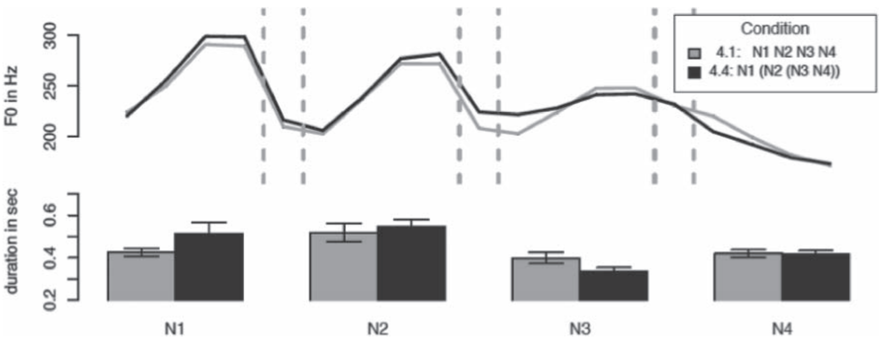


Fig. 4: Comparison of embedded right-branching condition (black) with baseline (grey).

first element of a grouping in these conditions and is thus compressed in pitch. A similar enhancement of downstep due to tonal compression was observed in the right-branching condition 3.3. In 4.2 and 4.4, the elements preceding a grouped constituent bear higher tones than the corresponding names of the baseline. Turning to duration, the baseline (grey) presents an unexpected pattern with N2 clearly longer, and N3 clearly shorter than N1. This durational effect is not accompanied by a similar effect in pitch. We will come back to this effect in the discussion (see Section 3.6 below). N1 of 4.2, a simplex element, is longer than in the baseline. Similarly, N1 of 4.4, which is in front of a left parenthesis, is also significantly lengthened, even more so than N1 of 4.2. This difference is explained by the fact that N1 in 4.2 is subject only to Similarity, whereas it is subject to both Anti-Proximity and Similarity in 4.4. In contrast, N3 in 4.2 and 4.4 are realized much shorter than in the baseline, but they do not significantly differ from each other (see also Figure 8 and Figure 9 for comparison). These are first elements of groupings and as such subject to Proximity. N2 is in both patterns located before a left parenthesis, but in 4.4, it is at the same time the first element of a recursive grouping. In the latter condition, it has a similar duration as in the baseline. Neutralization at the end of the sentence is once again observed in all patterns.

The left-branching structures in 4.3 (Figure 5) and 4.5 (Figure 6) differ from the baseline in several respects. Except if it is the last one in the sentence, the rightmost element of a grouping is higher in pitch than in the baseline. This explains why N2 in 4.3 and 4.5 as well as N3 in 4.5 are the highest points in these sentences. In all three patterns, N1, the first element of the groupings, is then realized at a lower level. The N2s do not present very large differences in their absolute values as compared to the baseline, but an upstep from N1 to N2 can be observed (whereas in the right-branching conditions, downstep was the rule). The duration relations of left-branching structures in 4.3 and 4.5 differ from the

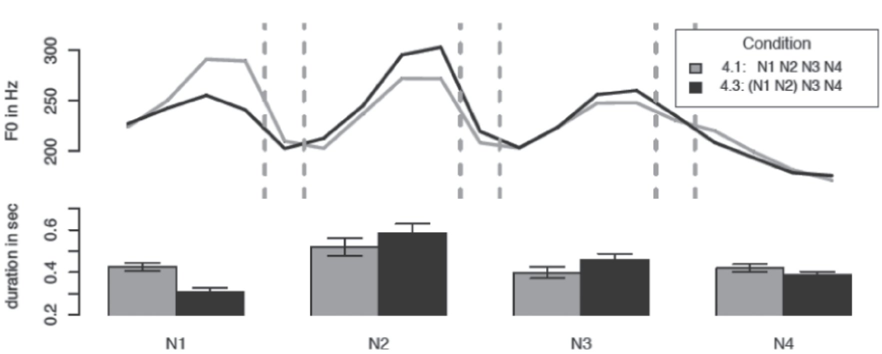


Fig. 5: Comparison of simple left-branching condition (black) with baseline (grey).

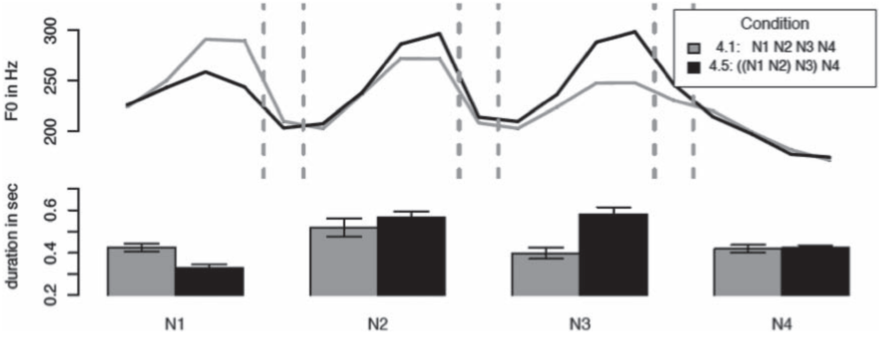


Fig. 6: Comparison of embedded left-branching condition (black) with baseline (grey).

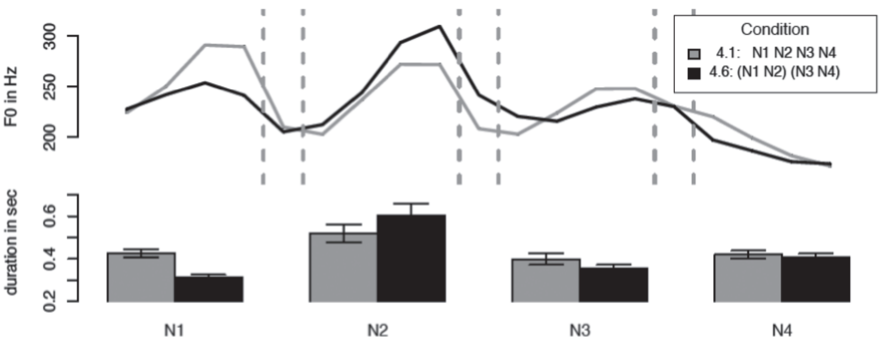


Fig. 7: Comparison of double grouping condition (black) with baseline (grey).

baseline in several respects. In both 4.3 and 4.5, N2 is located in front of a (non-final) right parenthesis. These names are significantly lengthened compared to the baseline. Moreover, N3 of 4.5, again preceding a right parenthesis, is longer than all other third names. In contrast, N1 in 4.3 and 4.5 is realized significantly shorter than in the baseline. Neutralization at the end of the sentence is once again observed in all patterns.

Finally, 4.6 with a double grouping is also compared to the baseline (Figure 7). In this pattern, we observe once more that the rightmost element of a grouping is higher and longer than in the baseline. This is the case for N2. N1, the first element of the grouping, is then shorter and is realized at a lower level, and an up-step from N1 to N2 can be observed. N3 is lower than in the baseline due to the fact that it is the first element of a grouping, and it is also shorter. As was observed in the three-name patterns, the downstep between N2 and N3 is larger than in the baseline.

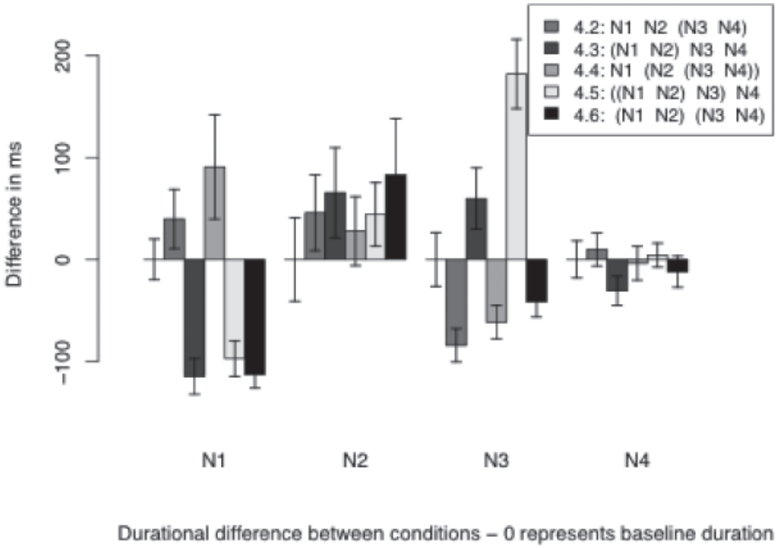


Fig. 8: Differences in duration between baseline and other conditions broken down by name.

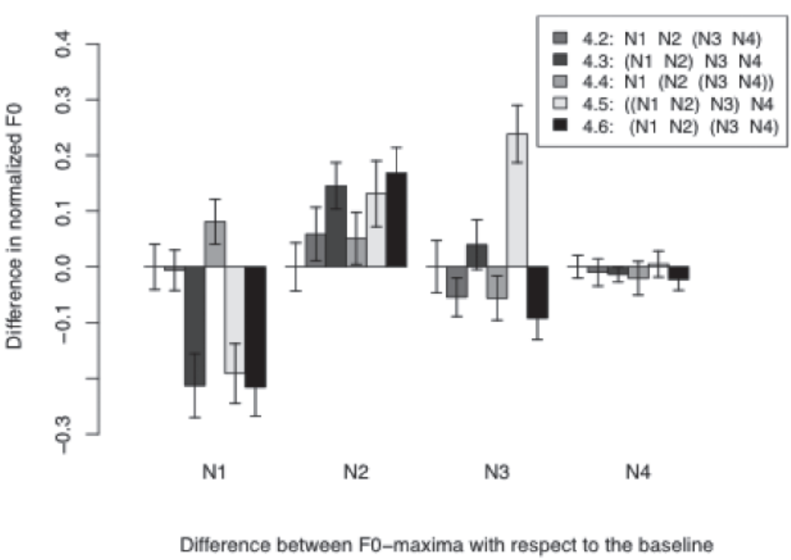


Fig. 9: Differences in normalized F0 between baseline and other conditions broken down by name.

Again, we generally find a strong correlation of duration and pitch.

As predicted, names that are subject to Proximity are shortened and compressed in pitch, while names that are subject to Anti-Proximity are lengthened and show upstep.

3.6 Discussion

In sum, the predictions of the Proximity/Similarity model are largely borne out. Each of the syntactic conditions appears to have a unique prosodic rendition, and the Proximity/Similarity model correctly predicts the prosodic effects that were observed: Names that are subject to Anti-Proximity are lengthened and show upstep, thereby strengthening a prosodic boundary. In contrast, names that are subject to Proximity are shorter and lower in pitch compared to the baseline, reflecting the cancellation of a prosodic boundary. The effect of Similarity appears to be weaker than that of Proximity or Anti-Proximity, but it still accounts for significant effects in duration (e.g. N1 of 4.1, N3 of 4.3).

A deviance in the parallelism regarding pitch and duration concerns the baselines 3.1 and 4.1. In 3.1, N2 was clearly longer than N1. Similarly, 4.1 displays a conspicuous lengthening of N2 and shortening of N3 compared to N1, but no comparable effect in pitch. According to the flat syntactic structure without grouping, the names were expected to be equivalent in duration across positions. We take the lengthening of N2 and shortening of N3 in the baseline 4.1 to be a reflection of abstract or 'inherent' grouping: even in the absence of syntactic motivation for grouping, speakers may favor a binary branching structure, which corresponds to an abstract grouping of N1 with N2 and N3 with N4. Independent evidence for such rhythmic grouping in the absence of explicit syntactic motivation comes from the prosodic rendering of telephone numbers: Baumann and Trouvain (2001) show that speakers preferably chunk a string of numbers into groups of two. Hunyadi (2006) reports a similar effect in a non-linguistic task: he presented Hungarian speaking participants with visual stimuli (4 equal-spaced dots in a row) and asked them to represent the visual display by mouse clicks. Measuring the time between clicks, Hunyadi found that participants needed more time between the second and third click than between the first and second. This effect of abstract grouping, however, was not confirmed in a speech production experiment in which participants read out a row of four letters. In any case, the tendency for abstract binary grouping without bracketing has a much weaker effect than the explicit boundaries in the binary branching structure of condition 4.6.

Overall, the right-branching structures (4.2 and 4.4) appear to be prosodically less articulate than the left-branching structures (4.3 and 4.5) and, correspondingly, right-branching structures are much more similar to the baseline. The prosodic markedness of the left-branching structures is considered to be due to the preponderance of upstep of boundary tones in these structures. Upstep is predicted for constituents that are subject to Anti-Proximity and is particularly strong if a non-final element that is subject to Anti-Proximity is preceded by an element that is subject to Proximity and thus compressed in pitch. The sequence of names which are subject to Proximity followed by names that are subject to Anti-Proximity is found in left-branching structures only. Correspondingly, the Proximity/Similarity model accounts for this specific prosodic markedness of left-branching structures as opposed to right-branching ones.

4 Model comparison

While the general predictions of Proximity and Similarity seem to be largely confirmed by the production data, we have yet to show how this model compares to other models of prosodic boundary likelihood or strength.

4.1 Method

In what follows, we evaluate the goodness of fit of the Proximity/Similarity model with the competing SBR and LRB algorithms. To do this, the boundary strength values that each theory predicts are calculated for each name of the structures 4.1 to 4.6.

For the Proximity/Similarity model, this is done as follows: The first factor Proximity has three levels: the baseline level is 0, i.e. all constituents of the baseline receive this predictor value. Names that are subject to Proximity are predicted to be shorter than the baseline; the corresponding predictor value is –1. For constituents that are subject to Anti-Proximity, the value 1 serves as the predictor. N2 in the right-branching condition with double embedding 4.4 is subject to both Proximity and Anti-Proximity. In this case, the two predictor values are simply summed, yielding 0 as the predictor for these constituents.

The second factor, Similarity, has two levels, 1 for names that are subject to Similarity and 0 for other names. The coding of the Proximity/Similarity model is summarized for the conditions with four names in Table 3.

Table 3: Coding scheme for the Proximity/Similarity model

Proximity/Similarity	N1	N2	N3
4.1 N1 or N2 or N3 or N4	Prox: 0, Sim: 0	P: 0, S: 0	P: 0, S: 0
4.2 N1 or N2 or (N3 and N4)	Prox: 0, Sim: 1	P: 1, S: 1	P: −1, S: 0
4.3 (N1 and N2) or N3 or N4	Prox: −1, Sim: 0	P: 1, S: 0	P: 0, S: 1
4.4 N1 or (N2 or (N3 and N4))	Prox: 1, Sim: 1	P: 0, S: 1	P: −1, S: 0
4.5 ((N1 and N2) or N3) or N4	Prox: −1, Sim: 0	P: 1, S: 0	P: 1, S: 1
4.6 (N1 and N2) or (N3 and N4)	Prox: −1, Sim: 0	P: 1, S: 0	P: −1, S: 0

Table 4: Coding scheme for the SBR model

SBR: boundary strength after	N1	N2	N3
4.1 N1 or N2 or N3 or N4	1	1	1
4.2 N1 or N2 or (N3 and N4)	2	2	1
4.3 (N1 and N2) or N3 or N4	1	2	2
4.4 N1 or (N2 or (N3 and N4))	3	2	1
4.5 ((N1 and N2) or N3) or N4	1	2	3
4.6 (N1 and N2) or (N3 and N4)	1	2	1

As for the SBR, the predictions are taken directly from Wagner (2005) and summarized in Table 4.

To adapt the LRB for our case, we need to make two deviations from Watson and Gibson's original algorithm: First, note that the LRB differs from the Proximity/Similarity model and the SBR in that it was designed to predict the likelihood of an intonational/intermediate phrase boundary in terms of the ToBI system (Beckman and Ayers, 1997) rather than the strength of a phrase break in terms of duration. Here, we consider the boundary strength to be reflected by the duration of the preceding constituent plus the following pause as dependent measure. As Wagner (2005) notes, "the advantage of this measure is that the annotation does not presuppose a theory of phrasing, and no labeling of prosodic categories (such as intonational phrase or intermediate phrase as in a ToBI-labeling) is necessary." Concurring with Wagner (2005), we will assume that the likelihood of an intonational/intermediate phrase boundary is strongly correlated to the duration of a prosodic break at any given position. In fact, Watson and Gibson themselves also use the term 'boundary weight,' which does justice to the gradient nature of prosodic boundaries. The second difference to Watson & Gibson's original approach is related to the nature of the materials used in the experiments. Compared to the sentences used in Watson and Gibson (2004), our

Table 5: Coding scheme for the adapted LRB model. Each predictor is the sum of the LHS (first addend), the RHS (second addend) and – where applicable – the addend 1 reflecting the end of the phonological phrase (cf. (13-c))

LRB: boundary likelihood after	N1	N2	N3
4.1 N1 or N2 or N3 or N4	1 + 1 = 2	1 + 1 = 2	1 + 1 = 2
4.2 N1 or N2 or (N3 and N4)	1 + 1 = 2	1 + 2 = 3	1 + 1 = 2
4.3 (N1 and N2) or N3 or N4	1 + 1 = 2	2 + 1 + 1 = 4	1 + 1 = 2
4.4 N1 or (N2 or (N3 and N4))	1 + 3 = 4	1 + 2 = 3	1 + 1 = 2
4.5 ((N1 and N2) or N3) or N4	1 + 1 = 2	2 + 1 + 1 = 4	3 + 1 + 1 = 5
4.6 (N1 and N2) or (N3 and N4)	1 + 1 = 2	2 + 2 + 1 = 5	1 + 1 = 2

structures are relatively short.[6] Therefore, IP boundaries are not necessarily expected. Correspondingly, we measure the complexity of the left-hand side and right-hand side in terms of phonological words rather than phonological phrases. At each word boundary, the boundary strength is calculated in accordance with (13) (cf. Watson & Gibson 2004).

(13) The LRB weight at a word boundary between w1 and w2 is defined to be the sum of
   a. the size of the left-hand side (LHS) constituent terminating at w1, measured in terms of phonological words (p-words);
   b. the projected size of the right-hand side (RHS) constituent in p-words starting at w2, if this is not an argument of w1;
   c. 1, if w1 marks the end of a phonological phrase.

The predictions of the modified LRB model are summarized in Table 5.

We compare the predictions of the Proximity/Similarity model with the predictions of the SBR and the LRB. Specifically, we evaluate the experimental results against the predictors of the three models. The duration of the individual items in each condition was averaged for each speaker. All models are mixed effects models that evaluate the log-transformed durations[7] of the names against the specific model predictors with speaker as random effect.

6 Watson and Gibson (2004) used sentences including relative clauses, such as *The director who the critics praised at a banquet insulted an actor from an action movie during an interview*.

7 log transformation is applied because the raw duration data is necessarily distributed in non-normal fashion, as there are only positive durations. Non-normal distribution would possibly violate the assumptions of the statistical model.

4.2 Results

Table 6 displays the modeling results for the Proximity/Similarity model. The formula in the upper row of each panel in Table 6 represents the linear model, which evaluates the dependent variable (logarithmized duration values) against the fixed effects (coded as described above). In the first model (upper panel), the single effect of the Proximity predictor (Prox) is evaluated; the second model evaluates the Similarity (Sim) predictor; in the third model (lower panel), the model estimates for the two fixed effects and the interaction are given. The variance that is due to the different speakers from the production experiment is accounted for in these models by including the variable "speaker" as a random effect term. As may be seen, the two fixed effects and the interaction account for significant portions of the distribution of the dependent variable (absolute t-values >2 indicate significance at $\alpha = 0.05$).

The SBR and LRB models are summarized in Table 7, which also displays a combined model with main effects of SBR and LRB plus the respective interaction. These three models confirm that LRB, SBR and the corresponding interaction have significant effects on the dependent variable.

That is, the predictors of all the models under consideration may each explain significant portions of the variance; however, we still need to determine which of the models (and which of the fixed factors) best explains the

Table 6: Parameters for models evaluating the Proximity factor (upper panel), the Similarity factor (middle panel), and the combined Proximity/Similarity factors and interaction

Formula:	**log(duration) ~ Prox + (1|speaker)**		
	Estimate	**Std. Error**	**t-value**
Prox	0.2742	0.00925	29.63
Formula:	**log(duration) ~ Sim + (1|speaker)**		
	Estimate	**Std. Error**	**t-value**
Sim	0.24108	0.02697	8.94
Formula:	**log(duration) ~ Prox × Sim + (1|speaker)**		
	Estimate	**Std. Error**	**t-value**
Prox	0.28928	0.01052	27.49
Sim	0.10853	0.01969	5.51
Prox:Sim	−0.16881	0.02683	−6.29

Table 7: Parameters for models evaluating the predictions of SBR (upper panel), of LRB (middle panel), and a combined model

Formula:	**log(duration) ~ SBR + (1|speaker)**		
	Estimate	**Std. Error**	**t-value**
SBR	0.26628	0.01520	17.51
Formula:	**log(duration) ~ LRB + (1|speaker)**		
	Estimate	**Std. Error**	**t-value**
LRB	0.16651	0.009532	17.47
Formula:	**log(duration) ~ LRB × SBR + (1|speaker)**		
	Estimate	**Std. Error**	**t-value**
LRB	0.36630	0.03499	10.469
SBR	0.52232	0.04895	10.670
LRB:SBR	−0.13394	0.0158	−8.478

variance in the data. To this end, a comparison of the fit of the models is in order.

As a measure of model fit, we take the R^2 value, i.e. the proportion of variability in the data set that the statistical model accounts for.[8] The R^2 values and the respective number of parameters (only fixed effects and interactions) are listed for each model under consideration in Table 8.

Evidently, the best model in terms of model fit is the Proximity/Similarity model, which clearly outperforms the combined SBR/LRB model. Note that both models make use of three fixed parameters (two main effects plus interaction term).[9] Therefore, the success of the Proximity/Similarity model is not simply due to the model's complexity. A model with Proximity as sole predictor fares second best, still outperforming the combined SBR/LRB model. However, the inclusion of Similarity is justified in that it significantly improves model fit, as determined by an analysis of variance comparing the simple Proximity model with a combined Proximity/Similarity model ($\chi^2 = 43.923$, df = 2, p < 0.001).

8 R^2 is the squared correlation of i) the fitted values of the model under consideration and ii) the actual duration values. R^2 can take values between 0 and 1 with 1 indicating a perfect fit.
9 All models also include the random effects parameter "speaker," so no difference in model fit is attributable to this parameter.

Table 8: Model comparison

Model	R^2	# of fixed effects
SBR	0.50	1
LRB	0.50	1
SBR × LRB	0.63	3
Sim	0.24	1
Prox	0.74	1
Prox × Sim	0.77	3

The success of the Proximity/Similarity model is probably due to the fact that it accounts for the weakening of a prosodic boundary between two names that are grouped together. Neither the SBR nor the LRB covers this effect. Instead, these models predict that the boundary after a left member of a grouped constituent is equivalent to the boundaries in the flat baseline structure.

All in all, the model comparison approach taken here suggests that the formulation of the Proximity/Similarity model has proven to be valuable. However, whether this model can account for the prosodic rendering of other syntactic environments is an open issue.

5 Perception experiment

As observed in the production experiment, the different syntactic groupings are reflected in different prosodic renderings.

The following perception experiment is conducted to answer the question whether listeners make use of the prosodic differences between the conditions, i.e. whether the appropriate syntactic structure is recoverable from the prosodic form. Specifically, we wanted to find out whether listeners recognize the syntactic structure that is determined by (recursive) syntactic embedding and the branching direction on the basis of prosodic information.

5.1 Predictions

The production experiment has revealed that each of the six syntactic conditions has a unique prosodic signature. Uniqueness of prosodic rendition, however, does not guarantee that the different conditions are easily discernable. How well the conditions can be recognized in perception depends for one thing on how strongly the conditions differ from each other in terms of prosodic rendition. Conditions that are marked by striking prosodic features are certainly more easily discernable compared to conditions that more closely resemble other conditions. That is, the higher the prosodic markedness, the better a certain syntactic structure may be recognized.

On the other hand, it may be more difficult for listeners to recognize syntactically complex structures, as these require higher processing costs. Accordingly, structures with recursive embedding should be more difficult to recognize than simply embedded structures.

Since the different left-branching structures (conditions 4.3 and 4.5) were marked by a very distinct upstep of boundary tones, it is hypothesized that these structures are more easily discernable than the right-branching structures (4.2 and 4.4), which all show a regular downstep pattern and more closely resemble the baseline pattern (condition 4.1).

Furthermore, we hypothesize that the conditions with recursive embedding (4.4 and 4.5) are more difficult to recognize than simply embedded structures (4.2, 4.3, and 4.6) or the baseline (4.1).

5.2 Methods

For each of the six conditions with four names, one sentence per speaker was arbitrarily chosen for the perception experiment. Correspondingly, the 21 speakers each contributed one sentence per condition (21 speakers × 6 conditions). The 126 resulting sentences were distributed over 3 blocks (each with 42 items) with speaker and conditions counterbalanced across blocks. In each block, the order of items was pseudo-randomized such that sentences of the same condition or the same speaker had a minimal distance of three items.

For each block, the individual sound files were pasted into a single sound string in the order determined by the randomization procedure. Each sentence was preceded by the auditory presentation of the sequence number spoken by the first author. The inter-stimulus interval was set to 4 seconds. The record level of the individual sounds was adjusted to 70db using an automated normalization procedure in praat. Forty-five listeners (15 per block) were equipped with an answer sheet and listened to the sequence of 42 experimental sentences over headphones. On the answer sheet, the six conditions were presented as abstract groupings with parentheses next to the corresponding item number. The format of the grouping is exemplified in (14) for condition 4.4.

(14) N1 (N2 (N3 N4))

While listening, the participants were asked to note on the answer sheet for each item which of the six conditions it belonged to by ticking the respective answer box. The presentation speed was determined by the recording. Listeners could not stop the presentation to listen again.

5.3 Results

Of the total 1890 presented items, 28 (1.5%) received no or no clearly identifiable response. These items were excluded from further analysis. For the 1862 (98.5%) valid responses, the confusion matrix in Table 9 shows the distribution with the presented condition tabulated against the condition chosen by the listeners.

Table 9: Confusion matrix tabulating the presented condition (rows) against the condition chosen by the listeners (columns)

	Chosen condition						total	Recognition precision
	4.1	4.2	4.3	4.4	4.5	4.6		
4.1 N1 N2 N3 N4	**260**	8	11	6	15	11	311	.84
4.2 N1 N2 (N3 N4)	15	**190**	16	20	23	44	308	.62
4.3 (N1 N2) N3 N4	5	14	**231**	11	33	17	311	.74
4.4 N1 (N2 (N3 N4))	10	127	20	**113**	10	28	308	.37
4.5 ((N1 N2) N3) N4	3	8	21	7	**264**	7	310	.85
4.6 (N1 N2) (N3 N4)	2	24	19	3	1	**265**	314	.84
total	295	371	318	160	346	372	1862	

The conditions were recognized overall with an accuracy of 71%, which is well above chance level (16.67%). The recognition precision for the presented conditions 4.1, 4.5 and 4.6 exceeds 80%; conditions 4.2 and 4.3 were recognized correctly less often (62% and 74% respectively).

As for the baseline 4.1 (84% recognition precision), the few misclassifications (n = 51) are relatively equally distributed across the competing conditions. The precision for the complex right-branching condition 4.4 is by far the lowest with only 37%. When presented with condition 4.4, listeners chose the simple right-branching structure 4.2 more often than the target structure (n = 127, 41%). That is, while listeners often recognized the branching direction correctly, they had problems identifying the depth of embedding in the right-branching structures. The confusion between 4.2 and 4.4 is asymmetric, however: if the simple right-branching structure 4.2 was presented, listeners correctly recognized it in 62% of the cases and most confusion occurred with condition 4.6 which was incorrectly chosen in 44 cases (14%). Note that, like 4.2, condition 4.6 also involves a grouping of the last two names.

Compared to the right-branching structures, the left-branching conditions 4.3 and 4.5 are not as prone to confusion with 74% and 85% correct classifications respectively. As for 4.3, most of the few incorrect classification answers concern condition 4.5; conversely, when listeners misclassified 4.5, they chose the simple left-branching structure 4.3 most often. That is, if listeners were presented with a left-branching structure (simple or complex) they recognized a left-branching structure in 88% of cases.

When presented with condition 4.6 (recognition precision 84%), most of the few misclassifications concerned the simple left-branching or the simple right-branching structure. Note that, just as 4.6, both 4.2 and 4.3 show strengthening of the prosodic boundary after N2; compared to N2, N3 is downstepped and significantly shorter in these conditions. This prosodic similarity might well explain the pattern of confusion.

For the statistical model, which evaluates the effects of syntactic embedding and branching direction on the recoverability of the structures, the following coding scheme was applied (see Table 10): For the first factor, syntactic embedding, the condition without embedding (baseline 4.1) was coded as 0, conditions with simple grouping (conditions 4.2, 4.3 and 4.6) were coded as 1 and conditions with multiple embedding (conditions 4.4 and 4.5) were coded as 2. For the second factor, branching direction, the left-branching conditions 4.3 and 4.5 were coded as 1, and the right-branching conditions 4.2 and 4.4 were coded as –1. Conditions 4.1 and 4.6, which lack a clear branching direction, were coded as 0.

A generalized linear mixed model (GLMM) with item, speaker and listener as random effects yields significant main effects for the fixed predictors embedding and branching direction as well as for the interaction. The results of this model, shown in Table 11, confirm that i) left-branching structures are more easily

Table 10: Coding scheme for evaluation of perception experiment

Condition	Embedding	Branch. Dir.
4.1 N1 or N2 or N3 or N4	0 (flat)	0 (neutral)
4.2 N1 or N2 or (N3 and N4)	1 (simple)	1 (right)
4.3 (N1 and N2) or N3 or N4	1 (simple)	−1 (left)
4.4 N1 or (N2 or (N3 and N4))	2 (double)	1 (right)
4.5 (N1 and N2) or N3) or N4	2 (double)	−1 (left)
4.6 (N1 and N2) or (N3 and N4)	1 (simple)	0 (neutral)

Table 11: Results of the GLMM on the perception data

	Estimate	Std. Error	z value	p value
Embed	−0.6156	0.1297	−4.747	<0.001
Branch	0.7420	0.3250	2.283	0.0224
Embed × Branch	−1.1200	0.2093	−5.351	<0.001

recognized than right-branching structures and ii) that increasing depth of embedding hampers recognition. The significant interaction reflects the fact that embedded left-branching structures are much less prone to confusion than embedded right-branching structures. Note that the doubly nested left-branching structure has the highest recognition precision of all conditions, while the doubly nested right-branching structure was recognized worst (cf. Table 9).

5.4 Discussion

As predicted, the left-branching conditions were better recognized than the right-branching conditions. Also, conditions with deeper embedding are more difficult to recognize than those with flatter structure, unless the former are clearly left-branching ones. The high recognition precision on the doubly nested, left-branching condition suggests that syntactic complexity does not hamper recognition if appropriate prosodic cues are provided. In contrast, the overall low precision on the right-branching structures reflects the shortage of adequate cues in these conditions.

Correspondingly, these results are best explained with recourse to the prosodic realization of the various conditions in the production experiment. The left-branching structures exhibit a distinct upstep pattern and clear pauses, which mark constituent boundaries. Such strong prosodic markedness is absent in the right-branching structures, which show regular downstep and thus resemble the baseline. As discussed above, upstep is particularly clear on a constituent that is subject to Anti-Proximity when it is preceded by a constituent that is subject to Proximity. We suggest that it is the specific upstep patterns and the corresponding boundary cues that make the left-branching structures easily recognizable. The depth of embedding has additional prosodic effects, namely the lengthening of simplex constituents in structures with grouped constituents (effect of Similarity). Although significant, this effect turned out to be rather weak in production and it might therefore only have had little effect on recognition in the perception experiment.

6 General discussion

6.1 The effects of Proximity and Similarity

Our experiments confirm that speakers use prosody for the rendition of syntactic grouping and embedding of coordinated names, thus disambiguating otherwise ambiguous structures. Conversely, listeners use prosody to retrieve the configuration intended by the speaker.

The two principles, Proximity and Similarity, account for the specific prosodic structure of the various grouping conditions in our experiment. The first principle, Proximity, accounts for the lower pitch and shorter duration observed on the left member of groupings compared to the flat structure of the baseline. Anti-Proximity has the opposite effect and strengthens the boundary between two constituents not grouped together. Such a boundary is expressed by longer duration and a greater hight of the high boundary tone. The second principle, Similarity, accounts for the observation that simplex elements in an expression containing groupings are lengthened. Arguably, this increased duration of simplex elements serves to achieve similar prosody to complex elements at the same level of embedding. The two principles guarantee that both branching direction and the depth of embedding have prosodic correlates.

A comparison of the Proximity/Similarity model with other models of prosodic boundary strength attests the P/S model's predictive power, at least for the structures tested in this experiment. The model comparison also reveals that the Proximity principle accounts for a much greater portion of the variance compared to the Similarity factor.

Although all conditions under scrutiny are distinguishable by virtue of prosody, the results show that prosodic cues are distributed asymmetrically: while right-branching structures are more similar to the flat baseline, left-branching structures are marked extensively by upstep and pauses at grouping boundaries. Accordingly, left-branching structures are more easily discernable in perception and significantly less prone to confusion than right-branching structures.

6.2 Recursion in prosodic structure

Recursion is understood as the embedding of grammatical constituents in constituents of the same kind. A sentence can be embedded in another sentence, or a noun phrase in another noun phrase. This property is uncontroversial for the syntactic structure of most languages. Traditional accounts of prosodic phonology explicitly deny that the same is true of prosodic structures, and the Strict Layer

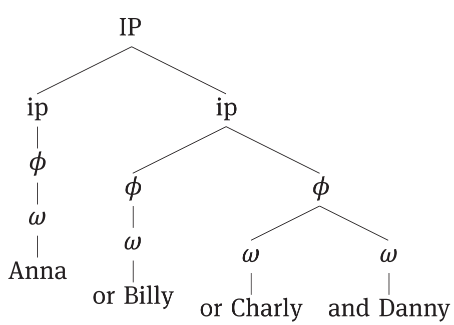
Hypothesis (SLH) of Selkirk (1984) and Nespor and Vogel (2007 [1986]) forbids recursion in prosody. In such a model, prosodic constituents can only iterate, that is, constituents of the same level can appear in a row but they cannot be organized hierarchically.

Based on the results of the production experiment, we claim that recursion in prosodic phrasing is a necessity if we do not want to allow uncontrolled profusion of additional prosodic levels.

The fine gradation of prosodic boundary strength, which systematically reflects the branching direction and the level of embedding, makes it difficult to interpret the results in terms of a strictly layered prosodic hierarchy that disallows recursion. Especially problematic is the ban on merging unlike prosodic categories, which the SLH imposes. If we conform to the SLH, in order to represent the prosody of a doubly nested coordinated NP made up of simple names (conditions 4.4 and 4.5 of the experiment), at least 4 prosodic categories are necessary. For demonstration, we may use the widely adopted categories ω (phonological word), ϕ (phonological phrase), ip (intermediate phrase) and IP (intonational phrase). Assuming that the IP, which wraps the complex NP, is part of a sentence and thus embedded within a larger prosodic domain, at least one additional larger prosodic category is needed. There is, however, no obvious category which could do this job – at least none for which there is independent evidence.[10] Therefore, the consequence of the ban on recursion is the uncontrolled and undesired profusion of stipulated prosodic categories.

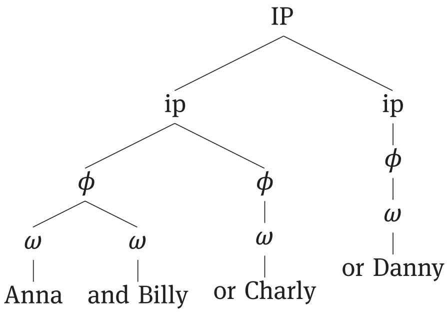
Moreover, according to the SLH, the first name in (15) would be equivalent to an intermediate phrase, even though it comprises only two syllables. The tension between the shortness of the name and its high status in the prosodic hierarchy is certainly counter-intuitive.

(15)



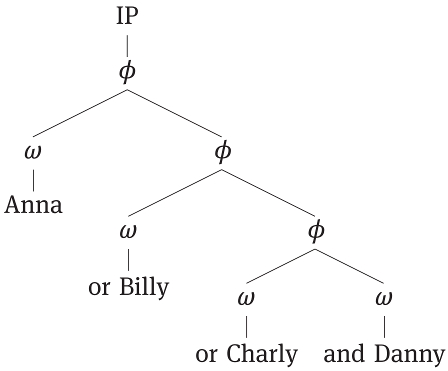
10 Clearly the 'Clitic Group' proposed by Nespor & Vogel (2007 [1986]) is not an adequate prosodic domain in this context. The proper names comprise at least a prosodic foot and thus cannot be subject to cliticization.

(16)

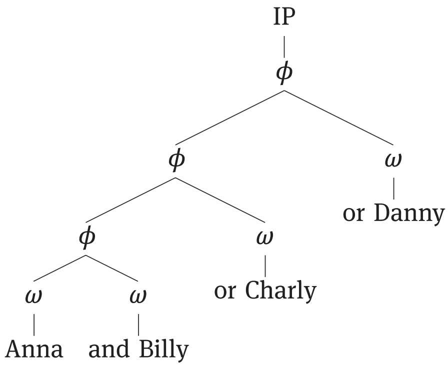


An alternative approach, which is in line with proposals by Ito and Mester (2012), Ladd (1986, 2008 [1996]), and Wagner (2005), explicitly allows recursion in prosodic structure. Recursively embedded syntactic NPs may thus be rendered as recursively embedded prosodic phrases. The device of recursion allows the generation of hierarchically ordered prosodic layers, without assuming different prosodic categories for each nesting level (cf. Ito and Mester, 2012). Also, in contrast to the SLH, prosodic constituents of different categories may be adjoined to form a prosodic constituent of a higher level. We assume that, in our case, each name corresponds to a prosodic word and grouped constituents form p-phrases of a higher order. The root node (or maximal prosodic projection) is represented as an intonational phrase. That way, the prosodic structure of the doubly embedded conditions can be represented much more economically (cf. (17), (18)).

(17)



(18)



An approach allowing recursion and merging of unlike prosodic categories predicts the prosodic differences between left-branching and right-branching structures that were attested in the experiment – differences that are not predicted within the SLH approach. Consider the representations that conform to the SLH. For both the right-branching (15) and the left-branching structure (16), the SLH predicts one ip-boundary, one ϕ-boundary and one ω-boundary between the four names (albeit in different orders); this would suggest that the prosodic structures should be equally complex – irrespective of the branching direction. In contrast, the recursive representation rightly predicts a difference in prosodic complexity between the two conditions: while (17) features no internal right boundary of a ϕ-phrase, (18) features two right edges of ϕ (after the 2nd and the 3rd name, respectively); in line with this representation, the left-branching structure proved to be prosodically more articulate in the experiment.

Given these considerations, we take our results to support the notion of recursion in prosodic structure. To sum up, we suggest that recursion of prosodic structure is clearly visible in German, and that speakers use it to disambiguate complex syntactic structure. The presence of prosodic recursion may be a feature of German (and other intonation languages), and does not need to be universal. Indeed, in an identical experiment with Hindi, reported in Féry and Kentner (2010), we showed that Hindi does not reveal the same prosodic features that have led us to assume recursion in German.[11]

11 An additional difference between German and Hindi is the robust head-final nature of Hindi as opposed to head-initiality in part of the syntax of German. It remains to be tested whether the 'articulate' prosody of German left-branching structures as opposed to the apparently inflexible prosody in Hindi is due to the difference with respect to head directionality between the two languages.

7 Conclusion

In this paper, we have shown the results of a production experiment with German speakers uttering sequences of three and four coordinated names, with different syntactic groupings. Our experiment was inspired by Wagner's (2005) work on English. The names were grouped in right- and left-branching structures, and two (of six) conditions for four names showed embedding of a group of names into a larger one. Groupings of names were always binary. A follow-up perception experiment was also performed in which other German speakers listened to the structures of the production experiment and had to decide which exact structure they had just heard. The results of both experiments were straightforward. German speakers and listeners heavily rely on prosody to disambiguate syntactic structure. Right-branching structures resemble the baseline, a sequence of names without any grouping, whereas left-branching patterns had different, more articulate realizations. Each single pattern had its own prosodic contour, although some patterns were more similar to each other than others.

We propose that the prosodic patterns are best accounted for by two principles called Proximity/Anti-Proximity and Similarity. Proximity claims that the default prosodic boundary separating each name from the next one is weakened when both names are grouped together. Anti-Proximity predicts strengthening of the boundary between two names that are not syntactic sisters. And Similarity requires that elements at the same level of syntactic embedding be separated by similar prosodic boundaries. While the Similarity component alone has relatively little predictive power, the Proximity/Similarity model as a whole is superior to both the Left hand side/Right hand side Boundary Hypothesis (LRB) of Watson and Gibson (2004) in which the size of the preceding and of the following syntactic constituents are the predictors for the likelihood of intonational phrase boundaries, and the Scopally Determined Boundary Rank (SBR) of Wagner (2005), which relates the strength of prosodic boundaries to syntactic levels of embedding rather than to the size of adjacent constituents.

As for the prosodic structure of German, the conclusion presenting itself is that recursion has to be assumed. The traditional Strict Layer Hypothesis (Selkirk 1984) cannot account for the kind of embedded structure exemplified in the paper. This confirms results of Féry and Schubö (2010) that showed the necessity of recursive prosodic structures in German.

References

Baumann, Stefan & Jürgen Trouvain. 2001. On the prosody of German telephone numbers. *Proceedings of Interspeech*, http://www.isca-speech.org/archive/ (24 January, 2013.)

Beckman, Mary E. & Gayle Ayers. 1997. *Guidelines for ToBI labelling*. The OSU Research Foundation.

Boersma, Paul & David Weenink. 2009. Praat: Doing phonetics by computer (version 5.1.12). *Software developed at the Institute of Phonetic Sciences, University of Amsterdam*.

Clifton, Charles, Katy Carlson & Lyn Frazier. 2002. Informative prosodic boundaries. *Language and Speech* 45. 87–114.

Clifton, Charles, Katy Carlson & Lyn Frazier. 2006. Tracking the what and why of speakers' choices: Prosodic boundaries and the length of constituents. *Psychonomic Bulletin & Review* 13. 854–861.

Cooper, William E. & Jeanne Paccia-Cooper. 1980. *Syntax and speech*, volume 3. Harvard University Press.

Ferreira, Fernanda. 1993. Creation of prosody during sentence production. *Psychological Review* 100. 233–253.

Féry, Caroline & Gerrit Kentner. 2010. The prosody of embedded coordinations in German and Hindi. In *Proceedings of Speech Prosody 5*, http://speechprosody2010.illinois.edu/papers/100014.pdf (24 January, 2013.)

Féry, Caroline & Frank Kügler. 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36. 680–703.

Féry, Caroline & Fabian Schubö. 2010. Hierarchical prosodic structures in the intonation of center-embedded relative clauses. *The Linguistic Review* 27. 293–317.

Féry, Caroline & Hubert Truckenbrodt. 2005. Sisterhood and tonal scaling. *Studia Linguistica* 59. 223–243.

Frazier, Lyn, Katy Carlson & Charles Clifton. 2006. Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences* 10. 244–249.

Gee, James P. & François Grosjean. 1983. Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology* 15. 411–458.

Grabe, Esther. 1998. Pitch accent realization in English and German. *Journal of Phonetics* 26. 129–143.

Hunyadi, László. 2006. Grouping, the cognitive basis of recursion in language. *Argumentum* 2. 67–114.

Ito, Junko & Armin Mester. 2012. Recursive prosodic phrasing in Japanese. In *Prosody matters*, Toni Borowsky, Shigeto Kawahara & Mariko Sugahara (eds.), 280–303. London: Equinox Publishing.

Kentner, Gerrit. 2007. Length, ordering preference and intonational phrasing: Evidence from pauses. *Proceedings of Interspeech* 8. 1385–1388.

Ladd, D. Robert. 1986. Intonational phrasing: The case for recursive prosodic structure. *Phonology* 3. 311–340.

Ladd, D. Robert. 1992. Compound prosodic domains. *Occasional Papers from the Linguistics Department*.

Ladd, D. Robert. 2008 [1996]. *Intonational phonology*. Cambridge University Press.

Lehiste, Ilse. 1983. The many linguistic functions of duration. *The Journal of the Acoustical Society of America* 73. S65.

Lerdahl, Fred & Ray Jackendoff. 1983. *A generative theory of tonal music*. The MIT Press.
Nespor, Marina & Irene Vogel. 2007 [1986]. *Prosodic phonology*. Mouton de Gruyter.
Pierrehumbert, Janet. 1980. *The phonology and phonetics of English intonation*. Cambridge, MA: MIT dissertation.
Pierrehumbert, Janet & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In *Intentions in communication*, Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), 271–312. The MIT Press.
Price, Patti, Mari Ostendorf, Stefanie Shattuck-Hufnagel & Cynthia Fong. 1991. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90 (6). 2956–2970.
Schubö, Fabian. 2010. Recursion and prosodic structure in German complex sentences. *Unpublished MA Thesis*.
Selkirk, Elisabeth. 1984. *Phonology and syntax*. Cambridge, MA: MIT Press.
Truckenbrodt, Hubert. 2002. Upstep and embedded register levels. *Phonology* 19. 77–120.
Wagner, Michael. 2005. *Prosody and recursion*. Cambridge, MA: MIT dissertation.
Watson, Duane & Edward Gibson. 2004. The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* 19. 713–755.
Wertheimer, Max. 1938. Laws of organization in perceptual forms. In *A source book of gestalt psychology*, Willis D. Ellis (ed.), 71–88. London: Routledge & Kegan Paul.
Wightman, Colin, Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91 (3). 1707–1717.

Length, Ordering Preference and Intonational Phrasing: Evidence from Pauses

Gerrit Kentner

Institut für Linguistik, Universität Potsdam, Germany
gerrit@ling.uni-potsdam.de

Abstract

This paper reports a speech production experiment in which the effects of surrounding phrase lengths and head-argument distance on intra-sentential pause duration were tested. While the results confirm an effect of phrase length on pausing, this effect is found to be distinctly stronger for long phrases preceding the pause than for long upcoming phrases. The results are discussed with respect to intonational phrasing tendencies and ordering preferences for unequal-sized constituents.

Index Terms: intonational phrasing, pausing, phrase length, phrasal ordering preferences, speech production.

1. Introduction

This study is concerned with intra-sentential pauses in read speech. The purpose of this study is to ascertain the effects of the length of preceding and upcoming material and its syntactic structure on pause duration at a given position within a sentence.

Pauses are an important feature of intonational phrase (IP) boundaries in speech [1, 2]. They are, however, a highly variable phenomenon governed by numerous factors such as speaking style, speaking rate, and the speaker's emphasis. One factor determining pause distribution and pause duration that has attracted the interest of several researchers [3, 4, 5] is the length of material preceding or following a pause. As Watson and Gibson [4] note, intonational phrase boundaries, and correspondingly pauses, can serve as refractory periods which are needed after the production of particularly long constituents. Similarly, a long upcoming constituent requires planning time which in turn might induce a pause. Watson and Gibson [4] formulate the LRB algorithm to predict IP boundaries at a given position within a sentence. It is mainly based on the number of words preceding and following that point such that the probability of an IP boundary rises when the preceding and/ or upcoming constituent is long. It is further constrained by syntactic factors. It has been shown that the likelihood of a pause and pause duration at a given position in a sentence increase with the number of syntactic brackets associated with that position [6]. That is, speakers preferably place phrase boundaries between rather than within syntactic constituents, i.e. speakers intonationally wrap the constituents [7].

A further constraint on IP boundary placement is again related to prosody: If possible, speakers concatenate IPs of approxiamtely equal size [8]. If the syntactic constituents to be concatenated in production have greatly differing lengths, the above constraints on intonational boundary placement conflict. In such a situation, speakers of English and German tend use the order short before long [9, 10]. The preference to place long IPs at the end of utterances can be seen as a prosodic constraint that might override syntactic requirements on constituent order, as evidenced by the phenomenon of Heavy NP Shift [11]. However, speakers do not always have the choice to order the constituents according to the needs of such prosodic constraints and might be forced to utter long constituents first.

Two recent studies on pause duration, namely [3] and [5] have scrutinised inter-sentential pause duration as an effect of preceding and following phrase lengths. In their study, Zvonik and Cummins [5] used synchronously read speech. The authors report that inter-sentential pauses shorter than 300ms almost exclusively occur when the preceding and following phrase consists of 10 syllables or less. The probability of a pause being short was shown to rise greatly if both the preceding and following phrase contained only 10 or fewer syllables, suggesting that the two predictors act superadditively.

Krivokapic [3] also used the method of synchronous reading. In her study on inter-sentential pauses, she compared pause length in four conditions, namely short/short, short/long, long/short, and long/long. Krivokapic found significant length effects for both preboundary and postboundary phrases, indicating that, irrespective of the order of the phrases, long phrases induce increased pause duration relative to short ones.

In a recent article, Watson and Gibson [12] have tested the hypothesis that the likelihood for an IP boundary increases with the integration distance between heads and their arguments. In their production experiment, however, they could not find a significant effect.

The present experiment is designed to ascertain the effect of the respective ordering of a long and a short constituent on pause duration between the constituents within German sentences. The syntactic structure of the experimental sentences is systematically varied in order to disentangle the different effects of syntax (i.e. head-argument distance) and constituent length on pause duration.

2. Methods

2.1. Experimental design

A speech production experiment is designed to test the influence of three factors on pause duration in speech production. These are 1) the respective ordering of a short and a long constituent, 2) the direction of the lexical head of the long constituent and 3) the position of the main verb within the sentence. With factors 2) and 3), the integration distance between syntactic heads and their arguments is systematically varied. All factors are two sided and crossed for this experiment.

Reading material is constructed according to the above factors resulting in 8 conditions. Sentences including coordinations are chosen as a test bed for this experiment. These constructions contain a proper name as subject in first position followed by a transitive verb frame with two coordinated objects. One of the object NPs is relatively long (10-15 syllables) and the other short (2-3 syllables). The lexical head of the long object is either preceded by a modifying adjective phrase (head right) or followed by a modifying PP (head left). The transitive verb frame either features with the main verb in second position followed by the objects or the main verb appears sentence finally and a modal verb occupies the second position. The respective structures are exemplified in (1).

(1)
a. [Subj] [Verb] [Obj1] [&] [Obj2]
b. [Subj] [Mod] [Obj1] [&] [Obj2] [Verb]

The stimuli for the production experiment consist of 24 sets of 8 sentences each. The experimental sentences were allocated to the subjects in a latin square design such that each subject would see only one sentence from each of the 24 sets. That way, each subject was presented 24 sentences, three from each condition. The 24 experimental sentences were embedded in 54 filler sentences. This set of 78 items was fed into a DMDX presentation [13] and pseudo-randomised for each subject such that sentences of the same condition did not appear adjacently.

In (2), the eight conditions are shown. The range specified in the parentheses refers to the relative position of the conjunction in that condition. It is calculated dividing the number of syllables preceding the conjunction by the total number of syllables in that sentence. As can be seen, the range for short-long sentences is distinctly below 0.5 while the range for long-short versions is higher than 0.5. The total length of the sentences ranges from 18-24 syllables.

(2)
a. short-long, V2, Head left (.25-.41)
b. short-long, V2, Head right (.25-.41)
c. short-long, V-end, Head left (.23-.33)
d. short-long, V-end, Head right (.22-.31)
e. long-short, V2, Head left (.75-.85)
f. long-short, V2, Head right (.75-.86)
g. long-short, V-end, Head left (.61-.77)
h. long-short, V-end, Head right (.63-.77)

In (3) the first 4 conditions are exemplified. Conditions e-h are made up of the same material but the order of the underlined objects is reversed.

(3)

a.
Paul malt <u>den Fluss</u> und <u>das Haus von Melanies Großtante</u>.
Paul paints <u>the river</u> and <u>the house of Melanie's grand aunt</u>

b.
Paul malt <u>den Fluss</u> und <u>das winzige geklinkerte Gartenhaus</u>.
Paul paints <u>the river</u> and <u>the tiny clinker-bricked summer house</u>

c.
Paul will <u>den Fluss</u> und <u>das Haus von Melanies Großtante</u> malen.
Paul wants to paint <u>the river</u> and <u>the house of Melanie's grand aunt</u>

d.
Paul will <u>den Fluss</u> und <u>das winzige geklinkerte Gartenhaus</u> malen.
Paul wants to paint <u>the river</u> and the <u>tiny clinker-bricked summer house</u>

2.2. Subjects

16 undergraduate students (7 male, 9 female) from the University of Potsdam took part in the experiment. All are native speakers of German and naïve to the purpose of the experiment. They either received course credit or were paid.

2.3. Recordings

Recordings took place in an acoustically shielded room with an AT4033a audio-technica studio microphone. Each subject was seated in front of a 15'' computer screen with the microphone placed approximately 30cm from the subject's mouth. A keyboard was placed on a table within close reach of the subject. Recordings were made on a computer using the RecordVocal function of DMDX and a C-Media Wave soundcard at a sampling rate of 44.1 kHz with 16 bit resolution.

2.4. Procedure

After a short instruction and three practice items (not part of the experimental set) the first sentence was presented on the screen. In order to enhance reading fluency, subjects were asked to familiarise themselves with the sentence and to press the space bar key afterwards. On pressing the space bar, the screen blanked for 200ms until the sentence reappeared on the screen. At this point, the subject's task was to read the sentence aloud. After that, a new sentence appeared and the procedure was repeated. For each sentence, there was only one realisation by subject. No corrections were recorded in the case of hesitations or slips of the tongue.

2.5. Data analysis

The data of the 16 subjects contains numerous slips of the tongue or hesitations due to self corrections (8.6%); the affected sentences were discarded. Overall, 332 sentences were manually annotated using the TextGrid device of Praat acoustic speech analysis software [14]. Duration analyses of the sentences were carried out automatically with Praat scripts. The durations of the two conjuncts and the duration of the silent interval before the conjunction were measured.

3. Results

Figure 1 shows the duration of the silent interval between the conjuncts as a function of the ordering of the long and the short conjunct.

The mean pause duration is 87ms for long-short sentences and 51ms for short-long sentences. A linear mixed effects model [15] with the crossed fixed factors "ordering" (short-long vs. long-short), "headedness" and "position of main verb" was employed; "subjects" and "sentence" were included as random effects. The logarithm of pause duration was chosen as the dependent variable. This model yields a significant main effect for "ordering" on logarithmised pause duration (t=-2.11, df=324, p=0.036). All other main effects and interactions remain non-significant.

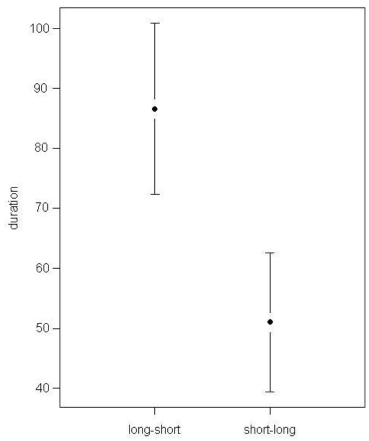


Figure 1: *Mean pause duration (in ms, with 95% confidence interval) at the conjunction as a function of the ordering of a long and a short conjunct.*

4. Discussion

These data confirm that surrounding phrase lengths have an effect on pause duration. Speakers pause longer at the conjunction of two unequal-sized conjuncts when the longer constituent precedes the shorter one. This result suggests that, in these asymmetric coordinations, the duration of the pause is positively correlated with the size of the preceding phrase but not with the size of the upcoming one.

Pauses are among the defining features of IP boundaries. The likelihood and strength of an intonational boundary grows with increasing pause length [1, 2]. The present evaluation of the experiment remains agnostic as to whether the silent intervals coincide with other IP boundary cues such as phrase-final lengthening or boundary tones. A closer examination of these cues would certainly be adequate to verify the results. However, given that items involving hesitation pauses were discarded, it is unlikely that the sentences under scrutiny contain pauses that interrupt intonational phrases. Since the intervals were measured at major constituent boundaries, namely at the conjunction, the dependent variable can be considered a good measure for intonational boundary strength. Therefore, it can be inferred from the result that the boundary strength and thus the likelihood of an IP boundary at the conjunction is higher in sentences with long-short ordering of the constituents than in short-long versions.

The findings of the present experiment complement and qualify Watson and Gibson's [4] LRB algorithm on IP boundary placement since only an effect of the length of the preceding phrase but not of the upcoming one can be confirmed. This is not to contest the results of Ferreira [16] and others who find that the size of an upcoming constituent is a predictor for pause length. However, it follows from the results here that the size of the preceding phrase is a stronger predictor for IP boundary placement. Watson and Gibson [4] themselves hypothesise that their LRB algorithm might be more successful when the relative influence of the upcoming phrase on boundary placement would be more restricted. This corresponds well with the notion of incrementality in the speech production process: A speaker does not always complete the planning of a constituent before he starts uttering it. Therefore, its ultimate size cannot be determined in advance and thus its influence on pause duration is limited.

The outcome of the experiment is especially interesting against the background of the short-long preference for constituent ordering in German and English [9, 10]. It seems that the violation of this preference has an effect on intonational phrasing. That is, while sentences which obey the preferred constituent order do not show a strong prosodic break, the long-short order tends to result in a more complex prosodic structure with an IP boundary between the unequal-sized constituents (as substantiated by the relatively longer pauses in this condition). A possible interpretation of this is that, when forced to utter the unpreferred order, speakers avoid a violation of the short-long preference on IP level by inserting an IP boundary after the long constituent. Thus, the IP can be seen as a domain for the short-long ordering preference.

Given the preference for long-short ordering of constituents in Japanese [17], it would be interesting to set up a similar experiment in that language to compare the effects of pause duration on preferred vs. dispreferred constituent order.

5. Acknowledgements

This research was financed by a grant from the German Research Foundation (DFG priority programme no. 1234) to the project "Prosody in Parsing" at the University of Potsdam. I would like to thank Caroline Féry, Frank Kügler, Anja Mietz, Umesh Patil, Shravan Vasishth, and three anonymous reviewers for helpful comments and suggestions.

6. References

[1] Cruttenden, A., Intonation, Cambridge University Press, Cambridge, 1997.
[2] Price, P.J., Ostendorf, S., Shattuck-Hufnagel, S., Fong, C., "The use of prosody in syntactic disambiguation", J. Acoust. Soc. Am. 90 (6), 1991, 2956-2970.
[3] Krivokapic, J., "Prosodic planning: Effects of phrasal length and complexity on pause duration", Journal of Phonetics, article in press.
[4] Watson, D., Gibson, E. "The relationship between intonational phrasing and syntactic structure in language production", Language and cognitive processes, 19, 2004, 713-755.
[5] Zvonik, E., Cummins, F., "The effect of surrounding phrase lengths on pause duration", Paper presented at Eurospeech 2003
[6] Gee, J.P., Grosjean, F., "Performance structures: A psycholinguistic and linguistic appraisal", Cognitive psychology, 15, 1983, 411-458.
[7] Truckenbrodt, H., "On the relation between syntactic phrases and phonological phrases", Linguistic Inquiry, 30, 1999, 219-255.
[8] Ferreira. F., Prosody, In: Nadel, L. (Ed.). Encyclopedia of cognitive science, Macmillan Reference, London, 2002.
[9] Wasow, T., "End-weight from the speaker's perspective", Journal of Psycholinguistic Research, 26, 1997, 347-361.
[10] Behaghel, O., "Von deutscher Wortstellung", Zeitschrift für Deutschkunde, 44, 1930, 81-89.
[11] Jackendoff, R.., "Foundations of Language". Oxford University Press, 2002.

[12] Watson, D., Gibson, E., “Intonational phrasing and constituency in language production and comprehension”, Studia linguistica, 59, 2005, 279-300
[13] Forster, K., Forster, J., “DMDX: A windows display program with millisecond accuracy”, Behavior Research Methods, Instruments and Computers 35, 2003, 116-124
[14] Boersma, P., Weenink, D. 2005. Praat: doing phonetics by computer. http://www.praat.org/ visited 5-Feb-07.
[15] Pinheiro, J., Bates, D., Mixed-Effects Model in S and S-PLUS. Springer, New York, 2000.
[16] Ferreira, F., “Effects of length and syntactic complexity on initiation times for prepared utterances”, Journal of Memory and Language, 30, 1991, 210-233.
[17] Yamashita, H., Chang, F., ”Long before short preference in the production of a head-final language”, Cognition, 2001, B45-B55.

DE GRUYTER MOUTON Zeitschrift für Sprachwissenschaft 2017; 36(2): 233–277

Gerrit Kentner*

On the emergence of reduplication in German morphophonology

https://doi.org/10.1515/zfs-2017-0010

Abstract: This paper discusses reduplication as a technique of word formation in German. In contrast to previous approaches, which consider reduplication as extra-grammatical and unproductive, this study identifies rhyme and ablaut reduplication as truly reduplicative processes in the morphology of German. A sizeable corpus of these reduplications and an acceptability rating study attest the productivity of this phenomenon. Other contemplable cases of reduplicative structures are properly treated as either phonological doubling, lexical sequencing, or (special cases of) compounding. An analysis in terms of Optimality Theory (OT) is offered which suggests that both rhyme and ablaut reduplication emerge when a segmentally and prosodically underspecified expressive morpheme is attached to a base – given that the base strictly obeys certain word prosodic requirements. The present approach considers the morphophonology to be blind to morphosyntactic structure and consequently eschews constraints that make explicit reference to base-reduplicant correspondence. The OT grammar successfully models the emergence of the fixed bipedal structure, the obligatory segmental deviance of the reduplicant, non-exponence of the expressive morpheme in the case of non-trochaic bases, the variable linearization of base and reduplicant in ablaut reduplication, and the interaction of reduplication with segmental alternations. Certain (crosslinguistic) correlations regarding constraints on reduplicative word formation and poetic devices, such as rhyme and meter, are discussed.

Keywords: ablaut, German, Optimality Theory, reduplication, rhyme

1 Introduction

This paper discusses reduplication as a technique of word formation in German. The central morphological devices for word formation in German, viz. derivation via affixation and compounding, generally concatenate segmentally speci-

***Corresponding author: Gerrit Kentner,** Institut für Linguistik, Johann Wolfgang Goethe-Universität Frankfurt, E-mail: gerrit@lingua.uni-frankfurt.de

fied bound or free morphs. The reduplicative exponent, in contrast, is a chameleon – its segmental realization is variable as it is largely determined by the stem to which it is attached. As an example, in the word *Hinkepinke* 'hopscotch' < *hinken* 'to hobble', the suffixed reduplicant[1] is disyllabic [pɪŋkə] while in the case of *Quitschquatsch* 'Fiddelsticks!' < *Quatsch* 'nonsense', the reduplicant is the prefixed monosyllable [kvɪtʃ]. The apparently flexible anchoring in addition to the variable segmental and syllabic structure of the reduplicant, and the playful and facetious connotation of these words suggest that this process is "extragrammatical" (Dressler 2000). In the same vein, Barz notes that reduplication in German "can hardly be dealt with systematically" (Barz 2015: 2407).

This treatise provides evidence to the contrary. Its objective is to demonstrate that reduplication in German not only deserves a systematic treatment, but also that it is possible to give an explicit formal account that captures the essential morphophonological features of reduplication. As opposed to previous work on reduplication in German (Bzdęga 1965; Schindler 1991; Wiese 1990), I suggest that the great diversity of seemingly or effectively reduplicating patterns, their use in mainly non-standard, spoken registers, and their unpredictability, have obscured the morphophonological regularity, and productivity, of two kinds of reduplication, namely rhyme reduplication (1) and ablaut reduplication (2).

(1)	*Schickimicki*	(< *schick*),	*Hasepase*	(< *Hase*),	*popelmopel*	(< *Popel*)
	trendy type	(< posh),	sweetheart	(< bunny),	nose picker (?)	(< bogy)

(2)	*Wirrwarr*	(< *wirr*),	*Mischmasch*	(< *misch*),	*Krimskrams*	(< *Krams*)
	jumble	(< woozy),	hotchpotch	(< mix),	bric-a-brac	(< stuff)

Alongside of rhyme and ablaut reduplication, several other kinds of reduplicative structures exist in German. In Section 2, I will propose and motivate a new taxonomy of the various structures, and discuss their status in the lexicon or grammar of German.

The morphophonological properties of rhyme and ablaut reduplication will be presented in detail in Section 3. It will be shown that these types of reduplication are closely related and subject to the same morphophonological constraints. The productivity of rhyme and ablaut reduplication in German is attested by a (most probably inexhaustive) corpus of previously undocumented

1 I use the terms *base* and *reduplicant* in a pre-theoretic, merely descriptive way: the base is the part of the reduplication most faithful to the stem's canonical realization while the reduplicant is the (partial) copy, which may deviate from the stem w.r.t. its segmental makeup; the identification of base and reduplicant is not possible in cases of identical reduplication.

cases (n > 150) of rhyme and ablaut reduplication (the corpus is listed in the appendix). Apart from the corpus, the acceptability of rhyme and ablaut reduplication is backed up by a questionnaire.

The sizeable number of these two types of reduplications, and their regular shape substantiate their morphological productivity. This motivates a grammatical analysis, formulated in terms of Optimality Theory (Section 4). This analysis suggests that rhyme and ablaut reduplication are in fact the result of quite regular morphological concatenation with the unusual proviso that the morpheme attached to the base be segmentally underspecified. It will be shown that the reduplicative nature of these words is due to purely phonological constraints. In Section 5, I discuss issues that the OT model does not resolve, and certain crosslinguistic correlations regarding constraints on reduplication and poetic devices, such as rhyme and meter. Section 6 concludes the paper.

2 Reduplicative structures in German

The diversity of reduplicative forms in German was documented by Bzdęga (1965) who presents an impressive collection of approximately 1880 reduplicative tokens gleaned from a broad range of dialectal and historical strata of German. However, only a subset of the patterns identified by Bzdęga (1965) are actually attested in modern Standard German, with only about 100 in general use (Wiese 1990). The two biggest classes of reduplicative structures in Bzdęga's collection are in fact forms with rhyme or ablaut, but many either lack a synchronically transparent morphological base (e.g. *Kuddelmuddel* 'mess'), or correspond to two stems (e.g. *Schnippschnapp* < *schnippen* 'to snip', *schnappen* 'to snatch') and are thus more properly treated as a special kind of compound (cf. Section 2.2). Other tokens are clearly reduplicative in nature but do not exhibit rhyme or ablaut (e.g. *Mama* 'Mum', *Kuckuck* 'cuckoo'). It thus appears that there is no uniform analysis available for the various patterns of reduplication found in German. Therefore, given the great diversity of reduplicative structures and their marginal status in morphological descriptions of German, a delimitation of the different types of reduplication is in order. To unveil the grammar of reduplication, it is necessary to isolate those patterns from the diverse set that are unambiguous instances of reduplication, and to assign the dubious cases a proper place in the lexicon and/or grammar of German.

As a first approach towards reduplicative structures in German, I propose a taxonomy of this diverse phenomenon. The taxonomy takes its cue from the degree of lexicality of the reduplicated form and the correspondence to an identifiable morphological base. It thus differs substantially from Bzdęga's classification, which is solely based upon the phonological makeup. The proposed classification is depicted in (3).[2] This taxonomy distinguishes i) syntactically autonomous, i.e. non-incorporated, interjections (3-I), ii) words used as lexical items (3-II), and iii) structures beyond the word, i.e. reduplicative phrases (3-III).

(3) Reduplicating structures
- I reduplicative interjections
  - (i) restricted to paralinguistic use, violating word phonotactic principles ([ts.ts.ts] – sound marking disapproval, *hahaha*, *hihi* – laughter, *rattattattatta* – imitation of machine gun)
  - (ii) phonotactically legal (onomatopoeic) interjections without lexical base (*dingdong* – imitation of doorbell, *piffpaff* – imitation of gun)
- II reduplicative forms used as lexical items
  - (i) no morphological base identifiable
    - a. purely phonological doubling, onomatopoetic words (*Mama*, *Kuckuck*)
    - b. synchronically unrecoverable base (*Techtelmechtel*)
  - (ii) with single morphological base: Rhyme and ablaut reduplication *Schickimicki* (1), *Krimskrams* (2)
  - (iii) combination of two stems
    - a. blends (*Schnippschnapp*), (recursive) compounds (*Kindeskind* 'grandchild', lit.: 'child of the child')
    - b. identical constituent compounds (*Reis-Reis* 'rice-rice')
- III reduplicative phrases
  - (i) frozen coordinations (*fix und foxi* 'to be done for')
  - (ii) X-and-X-construction (*teuer und teuer* 'expensive and expensive')
  - (iii) lexical sequence (*sehr sehr schön* 'very very nice', *schnell schnell* 'quick quick')

According to a widely accepted definition (Rubino 2011), reduplication in the strict sense is understood as a repetition of phonological material within a word for grammatical or semantic purposes. Generally, the base for reduplication is a segmentally specified string, and a morphologically complex word is formed by adding a copy of (some part of) the base (but see Inkelas and Zoll [2005];

2 Here, and in the following, I will disregard reduplications with linking elements (*klapperdiklapp* < *klappern*, *Edepopede* < *Ede*, proper name) and exceptional cases like triplication (*pipapo* < ?, *rirarutsch* < *rutsch* 'slide').

Inkelas [2008] for a different definition including doubling at the morphosyntactic level). From the diverse set of reduplicative structures in (3), only the type in (3-II-ii), viz. rhyme and ablaut reduplications, accords with Rubino's definition. All other kinds of structures either lack the status as word (3-I) and (3-III) or do not form morphologically complex words (3-II-i), or operate on more than a single base (3-II-iii).

The distinction between reduplication proper and the various other reduplicative structures will be made explicit in the following subsections. The class of interjections (3-I) will be disregarded.

2.1 Phonological doubling

The above definition of reduplication, in making explicit reference to a morphologically defined base, justifies a distinction between reduplication of an (or part of an) identifiable morph on the one hand and "phonological doubling" (4), where no morphological base can be identified.

(4) a. *Mama*, *Papa*, *Pipi*, *Kaka*
Mum, Dad, pee, poo-poo

b. *Tamtam*, *plemplem*, *ballaballa*
fuss, batty, crazy

Conceivably, the underlying segmental specification for the words in (4a) consists in just a consonant and a short vowel, e.g. [ma] in the case of *Mama*. In order to become a legitimate word, the segmental structure needs to be augmented to form at least a bimoraic foot; doubling might serve this need. In that sense, words like those in (4a) do not operate on, or add, morphological structure; instead, the whole word becomes the morphological stem that is then open to morphological processes such as inflection or compounding. According to Saba Kirchner (2010), the kind of purely phonological doubling might be characterized as a repair mechanism that becomes necessary when the segmental specification of the morpheme is too sparse to form a minimal word on its own.[3]

In contrast to (4a), the duplicated forms in (4b) do not suffer from moraic sparseness. However, since the base for the doubling process does not have a

3 Note that, in German, the unreduplicated forms *Ma* [ma:] and *Pa* [pa:] are not uncommon. They are, however, clearly perceived as anglicisms. Therefore, I assume that these forms are the result of hypocoristic truncation that is common in English (e.g. Lappe 2008; Alber and Arndt-Lappe 2012) rather than bases for reduplication themselves.

morphological status on its own, these cases are subsumed under the same rubric.

It is interesting to note that phonological doubling generally preserves segmental identity between base and copy while the prototypical cases of morphological reduplication, i.e. rhyme and ablaut reduplication (*Hinkepinke*, *Quitschquatsch*), generally foster non-identity of base and reduplicant. The issue of segmental (non)identity in reduplication will be discussed further in the Sections 2.3, 3, and 4.

2.2 Reduplication vs. compounding

If more than one base can be identified, the resulting word may be considered a (special kind of) compound or blend – irrespective of the segmental similarity between the bases (cf. Bzdęga 1965; Schindler 1991). This holds especially for paronomastic words like *Klimperwimper* ('person blinking one's eyes' < *klimpern* 'to tinkle' and *Wimper* 'eyelash') or *Schnippschnapp* ('snap' < *schnippe(l)n* 'to flick' and *schnappen* 'snap') that resemble rhyme or ablaut reduplications.[4] Often, however, in these cases, we cannot definitely exclude reduplication proper as the responsible process of word formation – it is at least conceivable that the reduplicant is accidentally homophonous to a stem that is not actually part of the word formation. As a case in point, consider the nickname *Sillepille*, derived from *Sille* (which in turn is a hypocoristic version of the female proper name *Silke*): the rhyming counterpart to *Sille* may be used as a noun (*Pille* 'pill' or 'ball'), but there is no obvious way to interpret *Sillepille* as a compound of the stems *Sille* and *Pille*; there is simply no semantic trace of *Pille* as denoting "ball" or "pill" in *Sillepille*. Moreover, according to the rules of compound formation in German, *Pille* would be assigned the role of head, and thus this compound would typically denote some kind of pill or ball. In contrast, *Sillepille* is exclusively used as a (facetious) reference to a person with the proper name *Sille*. I therefore treat cases of this kind as reduplications with segmental alternations (see below).

Krimskrams (< *Kram(s)* 'stuff') represents another problematic case: **Krim(s)* is most likely diachronically related to *Krümel* 'crumb'; however, there is no synchronically available stem **Krims* that would permit a transparent analysis as compound. I therefore treat *Krimskrams* as ablaut reduplication. A

4 The question of whether *Schnippschnapp* is a blend or a compound and whether blending is just a special form of compounding is beyond the scope of this paper. The reader is referred to the discussion in Ronneberger-Sibold (2006).

conceivable alternative would be to analyze **Krims* as a “cranberry morpheme” that attaches uniquely to *Krams*. This analysis however would beg the question of why the two stems are identical except for the ablauting vowel, a fact that is straightforwardly explained under the reduplication analysis (see Section 4). Several similar items exist in which only one of two diachronically identifiable stems are synchronically used.

A further case of doubtfulness concerns the verbs *schlampampen* < *schlampen* ‘to be sloppy’, *rumpumpeln* < *rumpeln* ‘to rumble’, *klimpimpern* < *klimpern* ‘to clink’. Wiese (1990) treats these cases as reduplications that come about due to the affixation of an underspecified syllable, segmental copying, and subsequent resyllabification. However, these verbs also allow for an analysis as blends where two phonologically similar stems are interleaved (cf. *schlampen + pampen*, *rumpeln + ?pumpe(l)n*, *klimpern + pimpern*). Whichever analysis is to be chosen as the correct one depends on the synchronic availability of the stems involved. In any case, this pattern does not appear to be productive and I am not aware of other tokens of this kind.

As these examples show, it is necessary to closely inspect the word under consideration in order to distinguish between reduplication and paronomastic compounding or blending. Whenever two synchronically transparent stems can be identified as constituting the word, the more conservative approach would suggest (a special form of) compounding – as long as there is no compelling morphosyntactic or semantic evidence against it (see above discussion on *Sillepille*).

2.2.1 Recursive compounds and identical constituent compounds (ICC)

Recursive compounds like *Kindeskind* ‘grandchild’ (lit.: ‘child of the child’), while presenting iterating phonological material, are not considered reduplications in the strict sense. The semantic transparency of these words suggests that they have to be treated as regular compounds made up of two identical stems. Note that the interpretation of *Kindeskind* is entirely analogous to other endocentric compounds like *Arbeiterkind* ‘working-class child’ (lit.: ‘child of a worker’). The linking element, which is often found in German compounds, is independent evidence for the compound analysis of these cases. Recursive compounds are restricted to only a few relational nouns like *Kind* ‘child’, *Helfer* ‘helper’, *Freund* ‘friend’.

As English and several other languages, modern Standard German exhibits identical constituent compounding (ICC) (Finkbeiner 2014; Freywald 2015; Hohenhaus 2004) a.k.a. contrastive focus reduplication (Ghomeshi et al. 2004).

Like recursive compounds, German ICC inherently involves doubling of a word, but ICCs are not confined to nouns (5a) but also used with adverbs (5b) or (predicative) adjectives (5c). Freywald (2015), who reports on a large-scale corpus search, did not find any verbs as bases for German ICC.

(5) a. *Nimmst Du Basmatireis oder einfach Reis-Reis?*
'Do you take basmati rice or just rice-rice (i.e. prototypical rice, standard variety rice)?'
b. *Was meinst Du mit 'jetzt' – jetzt-jetzt oder in zwei Minuten?*
'What do you mean by 'now' – now-now or in two minutes?'
c. *Der Typ ist echt schlau – nicht nur gewieft, sondern schlau-schlau.*
'This guy is really smart – not just slick, but smart-smart.'

As Ghomeshi's term 'contrastive focus reduplication' suggests, ICCs are used exclusively in contrastive contexts to denote the stem's prototypical features vis-à-vis less prototypical but contextually available alternatives. I will follow Hohenhaus (2004) and argue that ICC are best analyzed as a special form of endocentric compound: as in endocentric compounds, the first, accented, part restricts the meaning of the (identical) head – in this case by emphasizing the head's prototypical or ideal properties. Ghomeshi et al. (2004) explicitly discard the compound analysis as ICC may involve parts-of-speech not typically used in compounding.

In fact, it may be that this type of word formation is more promiscuous than canonical compounds w.r.t. to the stem that is used: ICC may target adverbs that are not typically used as stems in compounds. In contrast to canonical compounds, linking elements are banned in ICC. However, the promiscuity regarding the stems involved and the lack of linking elements are by no means compelling arguments against the compound analysis. Note that German makes productive use of phrasal compounds (Meibauer 2007), which generally lack linking elements. Furthermore, (phrasal) compounds may involve parts of speech in head or modifier position that are not typically found in canonical compounds (e.g. pronouns: *Über-Ich* 'superego', *Ich-AG* 'You Inc.', *Wir-Gefühl* 'group identity' or adverbs *im Hier und Jetzt* 'in the here and now').[5]

A cross-linguistic comparison buttresses the hypothesis that ICCs are properly treated as compounds: While German speakers find that the left part of an ICC modifies the meaning of the right part (the head), Italian and French speakers – to the extent that they use ICC in their native languages (Wierzbicka

5 Admittedly, in head position, those pronouns are certainly converted to nouns that may take a determiner.

2003) – identify the left part as head and the right part as modifier (Emanuela Sanfelici and Fatima Hamlaoui, p.c.). This correlates perfectly with the head-modifier ordering in canonical compounds in these languages. The head initial ordering of romance is also used in the Linnaean binomial nomenclature of biological taxonomy that makes use of latinate bases. As in romance ICC, identical names for biological genus (left member, head) and species (right member, modifier) likewise indicate the species' prototypicality (e.g. *Bubo bubo* denoting the "prototypical" – from a eurocentric viewpoint – Eurasian eagle owl, as opposed to e.g. the snowy owl *Bubo scandiacus*).

Furthermore, again in parallel to phrasal compounds but in contrast to rhyme and ablaut reduplication, ICC are generally not lexicalized – instead, they are created *ad hoc* as they are bound to a salient contrastive context in order to be used.[6] Furthermore, other than rhyme and ablaut reduplication, ICCs do not appear to be prosodically constrained (see Section 2.3).

On the basis of these arguments, both recursive compounds and ICC may be distinguished from proper reduplication and assigned a place in the realm of German compounding (cf. discussions in Finkbeiner 2014; Freywald 2015; Hohenhaus 2004).

2.3 Reduplication and stem correspondence beyond the word – lexicalization versus *ad hoc* construction

Two constructions with corresponding stems that are not part of the same word are related to reduplication and therefore deserve at least brief discussion here.

First, many frozen coordinations ("Paarformeln" Müller 1997; Cooper and Ross 1975) also involve two phonologically corresponding, yet non-identical, stems as conjuncts (6).

(6) a. *hegen und pflegen* 'to nourish and cherish'
schalten und walten 'to have carte blanche'
b. *fix und foxi* 'to be tuckered out'
dies und das 'this and that'

These constructions are generally lexicalized and idiomatic. Furthermore, these binomials appear to be prosodically constrained, that is, the conjuncts are usu-

6 Based on a questionnaire study, Finkbeiner (2014) notes that ICCs may be plausibly interpreted even without explicit information about the context. However, the participants in her experiment also showed a high degree of uncertainty as to how to interpret ICCs in isolation.

ally confined to the size of a prosodic foot. Many, but not all, of these frozen coordinations display a rhyme or ablaut relationship of the conjuncts involved.[7] In these respects, they resemble the canonical cases of reduplication in (1) and (2). In contrast to reduplication, however, the rhyming or ablauting conjuncts are generally both lexical words – but the meaning of the coordination phrases is not necessarily compositionally transparent.[8]

As a counterpart to frozen coordinations, the "X-and-X-construction" (7) is a coordination of two identical (lexical) words (Finkbeiner 2012, for German).

(7) X-and-X-construction
a. *Schade dass die so teuer sind.*
'It's a shame they are so expensive.'
b. *Naja*, **teuer und teuer** – *wenn die Qualität stimmt, dann finde ich den Preis okay*.
'Well, expensive and expensive – if the quality is good, the price is fine with me.'
(Finkbeiner 2012: 1)

The coordinative construction contributes a contrastive reading of the two instances of the same word. As the core lexical meaning of the two conjuncts is trivially identical, the contrast is only valid when the specific situational context and the relevant pragmatic meaning are taken into account. That is, the X-and-X-construction is used to "negotiate the situational meaning of a previously used lexical item" (Finkbeiner 2012: 1, and references therein). Interestingly, the X-and-X-construction is a syntactically and prosodically autonomous entity. It cannot form a syntactic argument unless in existential constructions. Correspondingly, X-and-X-constructions are intimately bound to a context in which the lexical item X has previously been used. As such, they are used productively and are generally *ad hoc* constructions that rarely become lexicalized.

Comparing, on the one hand, rhyme and ablaut reduplication with identical constituent compounding (ICC, cf. Section 2.2.1), and, on the other hand, frozen coordination with the X-and-X-construction, an interesting correlation arises concerning the segmental (non-)identity and the ability to become lexicalized (see Table 1). This correlation appears to hold on the word level and on

7 *Kaffee und Kuchen*, lit.: 'coffee and cake', 'five o' clock tea' makes use of alliteration to establish a phonological correspondence between the conjoined stems.

8 There are several frozen coordinations, in which one stem is synchronically not in use, e.g. *holtern und poltern*, lit.: 'to ? and to rumble.'

Table 1: Correlation of segmental identity and lexicalization in word-like and phrase-like reduplicative constructions.

	word level	**phrasal level**
lexicalized (segmental non-identity, prosodically constrained)	heterogeneous reduplication *Schickimicki*, *Mischmasch*	frozen coordination *hegen und pflegen* *fix und foxi*
ad hoc (segmental identity, prosodically unconstrained)	ICC *Ausbildung-Ausbildung* *Reis-Reis*	X-and-X-construction *gerecht und gerecht* *teuer und teuer*

the phrasal level: Both ICC and the X-and-X-constructions are characterized by segmental identity of the stems involved. These constructions are generally unconstrained by prosodic factors, that is, non-trochaic or polysyllabic bases can easily undergo ICC or may be used in X-and-X constructions.[9] At the same time, these nonalternating constructions are usually contextually bound *ad hoc* formations which rarely become lexicalized.

In contrast, rhyme/ablaut reduplication and frozen coordinations necessitate segmental alternation of the constituting feet/stems; they are generally morphologically simple lexical items or at least lexicalizable (the latter in spite of their phrasal nature) and, furthermore, prosodically constrained. At this point it is not clear to me what may cause this correlation of lexical status and segmental (non-)identity/prosodic shape, but since it holds on the word and on the phrasal level, it is most probably not accidental. However, the fact that ICC and X-and-Xconstructions are neither affected by segmental alternation nor by prosodic shape constraints indicates that their segmental material is fully specified in the input. This again lends credibility to the claim that ICCs are not reduplications in the strict sense, but (phrasal) compounds.

2.4 Lexical sequences

There are several cases in which the base is clearly a word (*hopp* ‘lollop, get a move on’) yet, in contrast to rhyme and ablaut reduplication, appears to reduplicate without alternation (*hopp hopp* ‘Get a move on!’).

9 Likewise, morphologically complex bases can undergo ICC or be used in X-and-X constructions.

(8) Lexical sequences[10]

dalli-dalli,	*hopp hopp*,	*los los*
hurry up,	get a move on,	go

The great majority of these cases has to be distinguished from proper reduplication as they fail to abide by the imperative of lexical integrity. Remember that, according to the above definition, the product of reduplication be a word. Schindler (1991) provides compelling reasons to question the lexical integrity of structures like *hopp hopp* or *dalli dalli*. Any sequence of identical morphs that is open to violations of lexical integrity (compare *dalli dalli* vs. *dalli*, *los*, *dalli*) qualifies as a lexical sequence rather than a word formed by means of reduplication in the strict sense. A concomitant of the separability is the possibility to unboundedly iterate the base, as in *sehr sehr sehr sehr schön* 'very very very very nice'. As Schindler (1991) notes, neither is characteristic of reduplication as a word formation technique (*Krimskrams* 'bric-a-brac' < *Kram(s)* 'stuff', **Dieser Krims*, *verdammt*, *(dieser) Krams*, **Krimskramskrims*). I will conclude that, whenever the lexical integrity of the reiterated forms are compromised, we can safely relegate these cases to the syntax.

3 Rhyme reduplication and ablaut reduplication

The two biggest classes of reduplication identified in the collection by Bzdęga (1965) are rhyme reduplication and ablaut reduplication. These are the only kinds of proper reduplication (as identified in the above section) that appear to be productive. An online search of these types of reduplication unearthed a wealth (> 150) of forms most of which have not previously been reported in the context of a treatise on reduplication let alone in a dictionary. Many but not all of these forms are derived from proper names; they may regularly be found as usernames in online platforms. Apparently, the obligation to create a unique username in internet forums leads to various kinds of formal augmentation (e.g. *vera123*, *Vera1982* < *Vera* etc.), and reduplication may serve the same purpose (*Veramera*), while adding a hypocoristic or facetious connotation. Consequently, when used as username or hypocoristic formation, reduplication results in nouns or, more specifically, proper names. In significantly rarer cases, bare verb stems may reduplicate, too. These may be used as proper names

10 The orthographic representation of these sequences is quite variable. *Dallidalli*, *dalli-dalli*, and *dalli dalli* are all attested.

(*Schwippschwapp* < *schwappen* 'to slosh', brand name for a lemonade), as interjections, e.g. in chat conversations (*mogelpogel* < *mogeln* 'to cheat'), but are also used frequently as modifiers within compounds where bare stems are commonplace (*Flitterflatter-Seidenband* < *flattern* 'flittering silk ribbon').

In general, apart from the use in online platforms, reduplications are commonly found in substandard registers of oral language, e.g. in playful conversation, not only with children. However, while rhyme and ablaut reduplications are mainly used in oral registers, they may also appear in press releases of big companies (cf. the following excerpt from a news item of a car manufacturer from 2002 in [9]). The fact that reduplications are used in such official contexts may be taken as a confirmation of the general intelligibility and acceptability of these word formations.

(9) *Knister-Knaster-Team bei AUDI spürt Geräusche auf*.[11]
'Crackle-RED-team at AUDI traces noises.'

The corpus currently consists of 94 rhyme reduplications and 61 ablaut reduplications that were collected by the author. These forms were either reported by informants, but also obtained via opportunistic searches in German-speaking internet domains. This list only features words with a single synchronically available base, so (paranomastic) compounds which, like *Klimperwimper* (< *klimpern* 'to tinker', *Wimper* 'eyelash'), are related to two bases, were excluded. Likewise, exceptional cases such as triplication and reduplication with linking elements were not collected. It has to be noted that the collection process was mostly blind to dialectal variation (but see discussion on $^{*/?}$*Michimachi* below), i.e. some of the tokens may be unacceptable in certain dialects. While the list is most probably not unbiased (as the author/collector is a native speaker of a northern variant of Standard German, his intuitions certainly affected the search for these forms), it may serve as a starting point for a more thorough investigation of this type of word formation. The examination of the corpus reveals a number of generalizations and several interesting correlations regarding the type of reduplication (rhyme vs ablaut) and other morphophonological properties of these words. These generalizations are listed in the following section.

11 http://www.presseportal.de/pm/6730/351204 [retrieved on May 20, 2016].
Note that *Knaster* 'weed, tobacco' is a lexical word in German. However, in *Knister-Knaster*, *Knaster* does not seem to hold any semantic relation to 'weed' or 'tobacco', and is therefore considered as reduplicant to *knister* 'to crackle'.

3.1 Morphophonological features of rhyme and ablaut reduplication

The examples in (10) and (11) illustrate the productivity of the two kinds of reduplication. Many of the forms listed are documented here for the first time.

(10) Rhyme reduplications
- a. suffixing reduplikation
  - (i) monosyllabic base
    Heinzpeinz, *Ralfpalf*, *Matzpatz*
  - (ii) disyllabic base
    Doppelmoppel, *Hasemase*, *Kallepalle*, *Popelmopel*, *Schorlemorle*, *Nickipicki*, *Michipichi*, *Rikepike*, *Silkepilke*
- b. prefixing reduplications not attested

(11) Ablaut reduplications
- a. suffixing reduplication
  - (i) monosyllabic base
    Mischmasch, *Ticktack*, *Wirrwarr*, *Stinkstonk*
  - (ii) disyllabic base
    Sillesalle, *Rikerake*, *Flügelflagel*, *kritzelkratzel*, *giggelgaggel*
- b. prefixing reduplications
  - (i) monosyllabic base
    Krimskrams, *Frinzfranz*, *Quitschquatsch*, *Zickzack*, *Mitzmatz*, *Mitschmatsch*, *schwippschwapp*, *Schnickschnack*, *pitschpatsch*, *plitschplatsch*
  - (ii) disyllabic base
    nigelnagel(neu), *schwibbelschwabbel*, *pipelpopel*, *rischelraschel*, *pickepacke(voll)*, *flitterflatter*, *krikelkrakel*

On the basis of the data in (10) and (11), several generalization w.r.t. the morphophonological behavior of rhyme and ablaut reduplication can be formulated:

(12) a. Reduplication results in strictly bipedal words[12]
b. The foot structure is strictly trochaic (bimoraic monosyllable or disyllabic trochee)

12 Reduplication with an unparsed linking element, e.g. $[[Ede]_\Sigma po[pede]_\Sigma]_\omega$ (reduplication of proper name *Ede*) are attested but I treat them as exceptional.

c. Base and reduplicant display the same number of syllables
d. The segmental makeup of base and reduplicant must not be fully identical → rhyme or ablaut
e. The sequencing of base and reduplicant is co-determined by phonological constraints in ablaut reduplication

These generalizations are in line with the fact that trisyllabic or quadrisyllabic bases (?*Nataliepatalie* < *Natalie*;[13] **Kunigundepunigunde* < *Kunigunde*, proper name) or iambic bases (**Ivonnepivonne* < *Ivonne* [ʔiˈvɔn]) cannot undergo reduplication without previous truncation to a trochaic foot via i-formation (*Ivipivi* < *Ivi* < *Ivonne*) (cf. Féry 1997; Grüter 2003; Itô and Mester 1997; Wiese 2001, on the grammar of i-truncations). Likewise, disyllabic words that superficially display a trochaic strong-weak syllabic pattern, yet consist of two feet, cannot become reduplicated. This ban holds for compounds (**Bahnhofpahnhof* < *[Bahn]*$_{\Sigma}$*[hof]*$_{\Sigma}$ 'train station') and for synchronically unanalysable yet prosodically complex proper names (**Gerhardperhard* < *[Ger]*$_{\Sigma}$*[hard]*$_{\Sigma}$, **Manfredpanfred* < *[Man]*$_{\Sigma}$*[fred]*$_{\Sigma}$).

3.1.1 Rhyme reduplication

The sequencing of base and reduplicant is fixed in rhyme reduplications but not in ablaut reduplications. In rhyme reduplications, the reduplicant invariably follows the base (see Table 2).[14]

The initial segment of the reduplicant is generally a labial, mostly [p], sometimes [m]. Koronal [d] is attested in loans from English (*okidoki*, *superduper*). The ban of segmental identity of base and reduplicant is reflected by the fact that bases with an initial labial invariably harness a different labial for the reduplicant (*Matzepatze* < *Matze*, **Matzematze; Pepemepe* < *Pepe*, **Pepepepe*). The question of why the suffixed reduplicant preferably starts in a labial sound cannot be answered with certainty. However, it is worth noting that, firstly, frozen co-ordinations exhibit the same tendency regarding the sequencing of

13 *Nataliepatalie* is only once attested on a German website; interestingly, it seems to be more commonly found on English-speaking websites although it is by far outnumbered by the reduplicated truncation *Nattipatti*. In general, it seems that English speakers find it easier to reduplicate trisyllabic bases. This fact may possibly buttress approaches to English metrical phonology that assume a greater variety of licit feet, e.g. including dactyls (Burzio 1994).

14 Compare this ordering to English reduplications like *helterskelter* < *skelte* (Old English) 'to hasten', *humblejumble* < *jumble* in which the meaningful base follows the reduplicant (Benczes 2012).

Table 2: Ordering of base and reduplicant, broken down by type of reduplication (rhyme versus ablaut). A chi-square test attests the non-independence of reduplication type (rhyme versus ablaut) and morph order ($\chi^2 = 61.42$, $p < 0.001$).

	Reduplication type	
Morph order	**Rhyme**	**Ablaut**
Base > Red (suffixing)	94	28
Red > Base (prefixing)	0	33

Table 3: Length of reduplication in number of syllables, broken down by type of reduplication (rhyme versus ablaut). A chi-square test attests the nonindependence of reduplication type (rhyme versus ablaut) and length in number of syllables ($\chi^2 = 25.97$, $p < 0.001$).

	Reduplication type	
No. of syllables	**Rhyme**	**Ablaut**
2 (monosyll. base)	5	24
4 (disyllabic base)	89	37

foot-initial segments (*hegen und pflegen* 'to nourish and cherish', *schalten und walten* 'to have carte blanche', *mit Sack und Pack* 'with bag and baggage') (Müller 1997) and this tendency appears to hold for other languages as well (Cooper and Ross 1975). Secondly, there is a general tendency for labials to occupy a foot-initial position (Torre 2003), and speakers may comply with this preference when given the occasion vis-à-vis a segmentally underspecified morph. The same tendency is reflected in children's and adult's productions when asked to provide a rhyme to a nonce word (Fikkert et al. 2005). Moreover, labials are used as onsets in several languages that make use of echo reduplication (Stolz 2008).

Rhyme reduplications exhibit a strong bias towards disyllabic trochees as the constituting feet (see Table 3). Monosyllabic bases are only rarely attested in the corpus (e.g. *Ralfpalf/Ralfmalf* < *Ralf*, *Heinzpeinz* < *Heinz*). These names seem to be more readily used in rhyme reduplication when augmented with the hypocoristic -*i*-suffix, resulting in disyllabic trochees *Ralfipalfi/Ralfimalfi* < *Ralf*, *Heinzipeinzi* < *Heinz*.

Furthermore, the corpus suggests that rhyme reduplications are avoided when the base presents with a complex onset (Table 4). In the rare cases of bases with a complex onset, the reduplicant invariably features a single onset consonant. Furthermore, apart from complexity, the onset of base and reduplicant are in most cases distinguished by the phonological feature, with the re-

Table 4: Onset complexity of the base, broken down by type of reduplication (rhyme versus ablaut). A chi-square test attests the non-independence of reduplication type (rhyme versus ablaut) and onset complexity ($\chi^2 = 14.61$, $p < 0.001$).

	Reduplication type	
Onset complexity	**Rhyme**	**Ablaut**
Singleton	85	39
Complex	9	22

duplicant presenting a single labial onset (*Fritzepitze*, ?*Fritzepritze* < *Fritz(e)*, *Klausipausi*, ?*Klausiplausi* < *Klaus(i)*).

The rhyme reduplications described here have to be distinguished from another case of reduplicative rhyming, viz. *shm*-reduplication (e.g. Nevins and Vaux 2003), that is found in Yiddish, in certain English registers and also, but only very marginally, in German (there are a couple of search engine hits for *Liebe Schmiebe* < *Liebe* 'love and such'). In contrast to rhyme and ablaut reduplication, *shm*reduplication may apply irrespective of the prosodic shape of the stem, at least in English (Saba Kirchner 2010).

3.1.2 Ablaut reduplication

The curious fact about ablaut reduplication is the variable ordering of base and reduplicant. Both prefixing and suffixing reduplicants are well attested (see Table 2 above). That is, any (morphological) constraint responsible for the ordering of base and reduplicant must be distinctly weaker than the phonological constraint regulating the ablaut order of the vowels [+high, +front] > [+low, +back]. In this respect, ablaut reduplication differs from other means of word formation in which the morph order is fixed.[15] While monosyllabic and disyllabic bases are equally attested (see Table 3 above), ablaut reduplication requires strict segmental restrictions w.r.t. the base in order to apply. Ablaut reduplication is prohibited if the stem vowel of the base cannot undergo ablaut, i.e. ablaut reduplication is restricted to bases with [i], [ɪ] or [a], [o], [ɔ].[16] Interesting-

15 In certain types of blends ("*Konturkreuzung*", Ronneberger-Sibold 2006), the fashion in which the morphs are interleaved or linearized is probably motivated by segmental or word prosodic features of the involved stems.

16 The only counterexample I am aware of is Christian Morgenstern's fictuous character *Flügelflagel* [flyːgəlflaːgəl] < *Flügel* 'wing'. Note however, that [y] is an acceptable off-rhyme to [i] (Primus 2002).

ly, in contrast to English ablaut reduplications (Minkova 2002) that invariably feature lax vowels, German ablaut reduplications also present tense vowels, although to a lesser extent (*nigelnagel* [niːgəlnaːgəl], *pipel-popel* [piːpəlpoːpəl]).

The requirement of ablautable vowels excludes bases with diphthongs and also those that exhibit more than one stem vowel (not counting schwa). Proper names like *Ina*, *Karlo*, *Tilo*, *Anna* have corresponding rhyme reduplications, but it is impossible to form ablaut reduplications like **Inaana*, **Kirlokarlo*, **Tilotalo*, **Innaanna*. Importantly, however, as an apparent exception to this generalization, stems that end in the i-suffix, which is productively used in truncations, can undergo ablaut reduplication *Mischimaschi* < ??*mischi* < *misch* 'mix', *Mitzimatzi* < *Matzi* < *Matthias*. That is, it appears that the hypocoristic i-suffix is invisible to this constraint on ablaut reduplication.

Ablaut reduplication appears to be blocked when it would entail further segmental alternations in the base or the reduplicant. A case in point is the [ç]-[x] alternation in Standard German, which demands that back vowels must not be followed by tautomorphemic [ç] (Wiese 1996). Consider the conceivable ablaut reduplication ?*Michimachi* < *Michi* < *Michael*. While attested in German internet domains, the places of discovery suggest that this reduplication is restricted to Upper German dialectal areas, likely those in which [x] may be used in the context of front vowels.[17]

The reason for the unacceptability of ?*Michimachi* in Standard German is most probably due to the alternation of the dorsal fricative triggered by the low back vowel ablaut [a] in the reduplicant, resulting in [mɪçimaxi]. While both feet are phonotactically well-formed, [mɪçimaxi] is not a licit reduplication in Standard German. This fact documents that base and reduplicant may only minimally diverge. A stronger correspondence between base and reduplicant would be warranted by the forms [miximaxi] or [miçimaçi], both of which are phonotactically illicit in Standard German. The same holds for the prefixing reduplication *Krichkrach* < *Krach* 'noise', which is attested in an Austrian source but does not seem to be felicitous in Standard German.

As in the case of rhyme reduplication, ablaut reduplication exhibits the same phonological tendencies w.r.t. morph sequencing as frozen coordinations, as witnessed by *dies und das* 'this and that'; *fix und foxi* 'to be tuckered out'. In fact, the ablaut ordering constraint requiring precedence of [i], [ɪ] over [a],

17 *Michimachi* and a dactylic variant *Michelemachele* are attested in the context of a sneering nursery rhyme with Upper German lexis (i) (*prunzt* for 'pee-3SG.PRS', *Kachi* for 'potty'), supporting the view that this specific token is restricted to this dialectal stratum:

(i) *Michimachi prunzt ins Kachi.*
'Michimachi pees into the potty.'

[o], [ɔ] appears to be universal (e.g. Cooper and Ross 1975). The ablaut order has been deemed a corollary of the general tendency to order short vowels before long ones (Minkova 2002), as the intrinsic length of the low vowels is generally greater than that of high front vowels (Lisker 1974).

> The essential empirical observation is that the stressed vowel in the left foot of these words is the shortest vowel in the system. Choosing the shortest possible vowel [ɪ] for the left peak of monosyllabic trochees would be the default case [...]. The stressed syllable of the right foot, on the other hand, fills its peak with the phonetically longest vowel available in the system which does not categorically violate the quantitative identity of the two parts. (Minkova 2002: 153)

Summarizing so far, the morphophonological regularity and productivity of rhyme and ablaut reduplication suggests that reduplication is synchronically available in the morphophonology of German. This does not imply that every speaker actively uses reduplication in everyday speech, the phenomenon remains one of the substandard (save a few forms that are in general use, e.g. *Mischmasch*, *Wirrwarr*, *Schickimicki*). However, it is predicted that forms that do not abide by the generalizations in (12) will be rejected by native speakers, while the ones that follow the generalizations should be acceptable. The following section reports on a questionnaire that puts this prediction to a test.

3.2 Experiment (questionnaire study)

In order to validate the generalizations stated in (12), an online questionnaire (SoSciSurvey, Leiner 2015) was devised in which participants judged the acceptability of various conceivable reduplicative structures. Specifically, participants judged prefixing and suffixing rhyme reduplications, prefixing and suffixing ablaut reduplications, full (or total) reduplications and reduplications of a single syllable. Moreover, the possibility of ablaut reduplication involving [ç]-[x]-alternation or ablaut reduplication of bases involving two full vowels was scrutinized. Finally, participants were asked to rate reduplication of non-trochaic bases.

3.2.1 Participants, materials and procedure

The questionnaire was announced in several introductory linguistics classes at Goethe University Frankfurt, usually taken by first year students of German. All in all, 72 students completed the questionnaire. Of these, 64 fulfilled the requirement of being native speakers of German (as per self report).

Participants were asked to judge whether a given base word could undergo reduplication when used in a playful context. Depending on the base, participants were given four, five or six reduplicative alternatives. Their task was to rate each of these on a five point Likert scale. The five points were assigned the following meanings:

5: I could use this word myself;
4: conceivable word, heard before;
3: conceivable word, never heard before;
2: virtually inconceivable that someone would use this word;
1: unacceptable as a word.

Five sets of items were devised. For the first set of items (set “r” for “rhyme”), four simple trochaic words were chosen: *Hansi* [proper name]; *Heinzi* [proper name]; *mogeln* ‘to cheat’; *hinken* ‘to hobble’. Apart from the suffixing rhyme reduplication (*Hansipansi*), participants were given a prefixing rhyme reduplication *Pansihansi*, the full or total reduplication *Hansihansi*, and a syllabic reduplication *Hahansi*.

With the second set of items (set “tri” for “trisyllabic”), the potential of (rhyme-)reduplication of four non-trochaic (amphibrachic) proper names (*Susanne*, *Tobias*, *Sabine*, *Elias*) was tested. Participants rated six reduplicative options: Rhyme reduplication (suffixing and prefixing: *Sabinepabine*, *Pabinesabine*), leftanchored and stress-anchored truncated rhyme reduplication *Sabipabi*, *Binepine*, full reduplication *Sabinesabine* and syllabic reduplication *Sasabine*.

The third item set (set “a” for “ablaut”) was devised to test the potential of ablaut reduplication: eight trochaic bases with stem vowel /ɪ/ (*Sille* [proper name]; *prickel* ‘tingle’) or /a/, /ɔ/ were chosen (*Gammel* ‘rot’; *Quatsch* ‘nonsense’; *zappel* ‘to fidget’; *Zottel* ‘dag’; *Matsch* ‘mud’; *Mops* ‘pug dog’). Here, participants were presented with ablaut reduplication (*Gimmelgammel*), the reverse ablaut reduplication (/a/, /ɔ/ > /ɪ/: *Gammelgimmel*), full (*Gammelgammel*), and syllabic reduplication (*Gagammel*).

A fourth item set (set “cx” for “[ç-x]-alternation”) was administered to probe the viability of ablaut reduplication when ablaut would induce [ç]-[x]-alternation. Four trochaic bases with /ɪ/ or /a/ followed by the dorsal fricative were chosen (*Michi* [mɪçi], *Achi* [axi], [proper names]; *Krach* [kʁax], ‘noise’; *sicher* [zɪçɒ], ‘sure’). Again, ablaut and reverse ablaut reduplication, full reduplication and syllabic reduplication were given as candidates to be evaluated.

The final item set (set “vv” for “two full vowels”) consisted of four trochaic bases with two full, potentially ablautable vowels (*Sara*, *Tilo*, *Jutta*, *Nina* [proper names]). Participants evaluated five options: reduplication with ablaut *Nina-*

nana and reverse ablaut *Nananina* of the head vowel as well as of the non-head vowel *Nininina*, *Ninanini* and full reduplication *Ninanina*.

The items of all five sets were presented in four differently randomized lists that were randomly assigned to participants. Every participant who answered the full item set was presented every item (24 items × 64 participants) and, for each item, judged four, five, or six reduplicative options. All in all, 6252 ratings were obtained.

3.2.2 Results and discussion

The following figures depict the ratings for i. the set of potentially rhyming reduplications (item set "r" and "tri"; Figure 1) and ii. the set of potentially ablauting reduplications (item sets "a", "cx", and "vv"; Figure 2). The diverging stacked bar charts (Heiberger and Holland 2015) represent the results of the Likert ratings as follows: the further to the right, and, concomitantly, the wider the lighter bar segments, the more acceptable the condition was rated. The two darkest shades represent the values "unacceptable" and "hardly acceptable". The three lighter shades represent "possible" or "conceivable" words with different degrees of familiarity or acceptability ("never heard – heard before – could use it myself"). The vertical reference line is set between the ratings representing "hardly acceptable" and "conceivable, never heard". In the following discussion of results, I will assume this line to represent the cut-off between acceptance and rejection.

As shown in Figure 1, suffixing rhyme reduplications of trochaic bases were deemed conceivable words in more than 80 % of responses, with nearly 50 % of cases for which raters reported to have heard or could use the words themselves. Prefixing rhyme reduplications, and full or total reduplications were clearly considered less acceptable (with rejection rates of ~50 %). Syllabic reduplications were rejected in ~80 % of responses. In the case of non-trochaic bases (set "tri"), acceptance rates for both types of truncated rhyme reduplications and the non-trochaic suffixing rhyme reduplication reached acceptance rates of close to 70 % (or were rejected in more than 30 % of cases). They were thus rated distinctly less acceptable than rhyme reduplication of trochaic bases. Prefixing rhyme reduplication, full reduplication and syllabic reduplication were rejected in the great majority of responses.

Similar results were obtained for ablaut reduplications based on monosyllabic stems or trochaic stems ending in a schwa syllable (set "a", Figure 2). Ablauting reduplications obeying the ablaut order of the vowels were rated as possible words in more than 80 % of cases. The reverse ablaut reduplication,

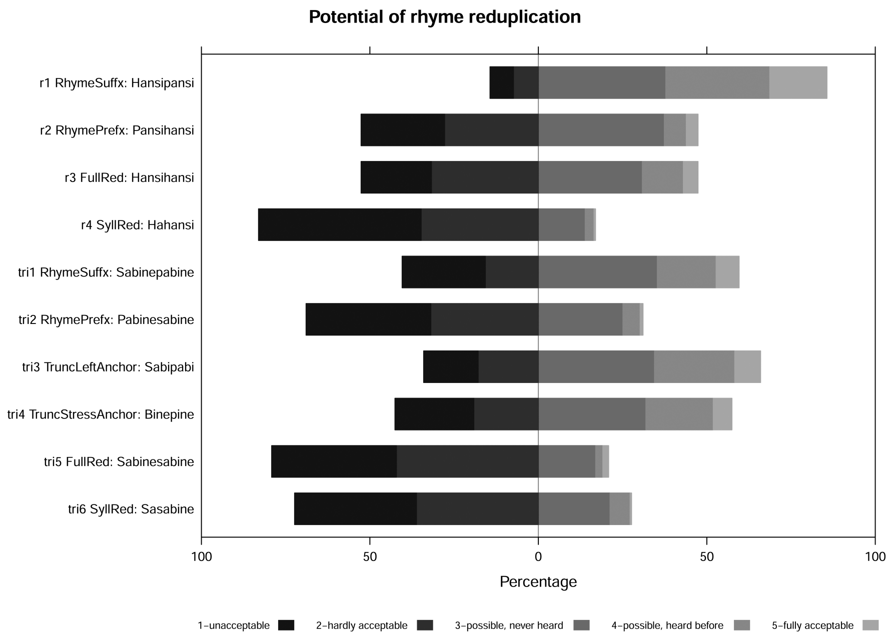


Figure 1: Ratings for rhyme reduplication and potential alternatives in item sets "r" (upper four bars) and "tri" (lower six bars). See main text for details.

full reduplication and syllabic reduplication were deemed acceptable in only 50 % of cases (full reduplication) or less (reverse ablaut and syllabic reduplication). The acceptance rate for ablaut reduplication is clearly reduced if ablaut induced the alternation of the dorsal fricative (set "cx", acceptance rate ~30 %) or when it was applied to bases with more than one full vowel (set "vv", showing acceptance rates of around 50 % for ablaut).

The results of the questionnaire largely confirm the predictions that result from the opportunistic word search reported above. Summarizing the findings, the great majority of responses indicate a general acceptance of rhyme and ablaut reduplications based on monosyllabic or trochaic bases. Ablaut reduplication is acceptable both for bases with /i/ and with /a/ as stem vowel, supporting the prediction that ablaut governs morph order. At the same time, the experiment yields corroborating evidence for the claim that ablaut reduplication is prohibited when it involves alternation of the dorsal fricative and when applied to bases with more than one stem vowel.

As expectable, the results of the experiment are graded rather than categorical. Not all participants were equally strict in rejecting candidates that the majority would deem unacceptable. However, the consistently high rejection

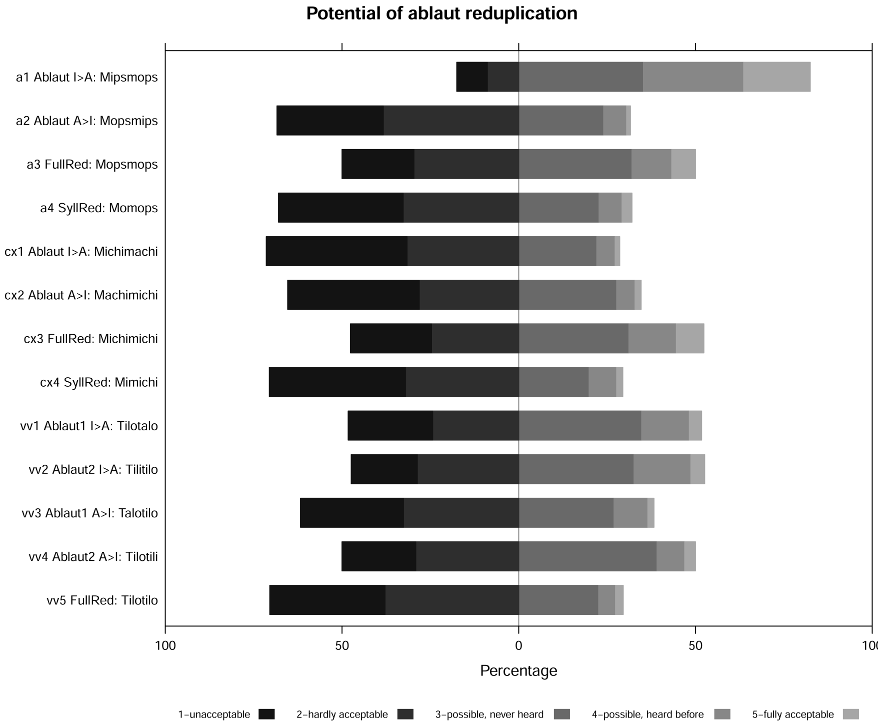


Figure 2: Ratings for ablaut reduplication and potential alternatives in item sets "a" (upper four bars), "cx" (5th to 8th bar from top), and "vv" (lower five bars). See main text for details.

rates for full reduplication, prefixing rhyme reduplication, reverse ablaut reduplication (/a/ > /i/) and syllabic reduplication suggest that, even though reduplication may be a phenomenon of playful substandard registers, there are clear boundaries to what is licit, and these boundaries may well be set by grammatical constraints that hold for the language at large. In the following section, I will put forward an OT-account sketching the grammar of rhyme and ablaut reduplication in German.

4 Rhyme and ablaut reduplication in OT

The optimality theoretic analysis entertained here follows the spirit of Bermúdez-Otero (2012) and Saba Kirchner (2010) who assume a strict division of labor between morphology and phonology, such that morphosyntactic information is not directly visible to the phonology module and vice versa. In this modular approach, the morphology's role is to assemble roots, stems, and affixes from the lexicon and concatenate them, thus determining a sequence of morphs complete with their lexically specified phonological content. The role of the phonology, then, is limited to interpreting this sequence of morphs without consideration of, or access to, the syntactic features or the morphological status (root, stem, or affix) of the underlying morphemes. Reduplication results when a segmentally underspecified prosodic constituent attaches to a base and needs to be spelled out. Provided that the empty morph cannot be filled with epenthetic material for independent reasons, it inherits segmental material from the base, leading to repetition of phonological material. In this latter respect, the present approach follows older ideas by Marantz (1982) and McCarthy and Prince (1995) et seq. who assume that reduplication is generally driven by the need to fill prosodic positions with segmental material.

However, in contrast to McCarthy and Prince's Base Reduplicant Correspondence Theory (BRCT, McCarthy and Prince 1995), the phonology in the present approach does not directly interact with, and remains blind to, morphosyntactic structure. Note that, according to McCarthy and Prince (1995), the phonology does not only regulate the mapping of underlying and surface structure; in addition, correspondence theory assumes parochial constraints for the mapping of input and reduplicant (IR-faithfulness), and, moreover, output-output correspondence, with faithfulness constraints evaluating the identity of base and reduplicant (BRfaithfulness). Equipped with this reduplication-specific machinery, the phonology has direct access to morphological information about which portion of the output corresponds to the base and which represents the reduplicant. Thus, BRCT assumes constraints that are specific to reduplication and are only relevant when an input morpheme is specified as being reduplicative (Red). Correspondingly, reduplication would be conceived of as a given morphological process rather than a phenomenon emerging within the phonology.

I will show here that rhyme and ablaut reduplication in German can be modeled without recourse to reduplication specific correspondence constraints. They can instead be modeled with purely phonological constraints that are anyway needed in the morphophonology of German and that are thus motivated independently.

In assuming a segmentally underspecified morph as the driver of reduplication, the present analysis contrasts with another take on reduplication, viz. morphological doubling theory (Inkelas and Zoll 2005; Inkelas 2008). Morphological doubling theory holds that reduplication is the result of concatenating two (segmentally specified) allomorphs, (each of) which may or may not be subject to certain construction-specific phonological processes (i.e. truncation). That is, reduplication is motivated in the morphology and as such, the repetition is already part of the input to the phonology; it does not serve phonological needs. Prime examples of morphological doubling in German were discussed above in section 2.2.1: recursive compounds and ICC, both of which have been shown to involve repetition at the morphosyntactic level and which do not adhere to the kinds of phonological conditions that affect rhyme and ablaut reduplication (i.e. segmental non-identity, fixed prosodic shape).

As discussed above, reduplication in German is mainly used for expressive purposes. Following Saba Kirchner's (2010) approach, I assume that reduplication results when a segmentally unspecified morpheme merges with a base. Given i) the strict footsized shape of the base and ii) the bipedal structure of German rhyme and ablaut reduplication, the expressive morpheme has to correspond to a foot (Σ) in these constructions (13). The foot, however, is underspecified w.r.t. its syllabic structure. The native lexicon of German mainly consists of monosyllabic or bisyllabic trochees (Eisenberg 1991); consequently, reduplications are either bisyllabic or quadrisyllabic.

(13) $\Sigma \leftrightarrow$ [+ expressive]

At the morphosyntactic level, the structure of rhyme and ablaut reduplication may be represented as the merger of a stem with an expressive head (14a). I assume the expressive morpheme itself to be morphosyntactically underspecified. As is often assumed for root compounds (e.g. Bauke 2014; Harley 2012; Roeper et al. 2002), a morpheme may take on various syntactic categories or it may remain a non-categorical root morpheme. This assumption fits well with the syntactic heterogeneity that is found in rhyme and ablaut reduplication: while many reduplications are hypocoristic or teasing forms of proper names (nominals), there are also verbal and adjectival reduplications, and also category-free instantiations used as modifiers in compounds (as in *Flitterflatter-Seidenband*, 'flittering silk ribbon').[18] The heterogeneity on the morphosyntactic side, however, by no means implies general irregularity. In fact, as described above and formalized below, the morphophonology is quite regular and may simply be assumed as a concatenation of the stem and the segmentally unspecified foot corresponding to the (syntactically likewise unspecified) expressive morpheme (14b).

18 Note that the word *schickimicki* can be either a (masculine) noun ('posh person') or a predicative adjective ('posh').

(14) a. Morphosyntax: expressive construction

Stem expressive

|| ||

b. Morphophonology: /stɛm/$_{\Sigma}$ + /Σ/

In the analysis to follow, I will specifically attend to the morphophonological level and its phonetic interpretation, ignoring for the most part the morphosyntactic features of reduplication.

Following the modular approach as expounded above, I will eschew constraints that are specifically geared towards reduplication. Instead, reduplication is shown to emerge from general phonological constraints that come into force whenever a segmentally underspecified morpheme is attached to a base. In line with Saba Kirchner (2010), the general scenario can be summarized as an interaction of the constraints in (15a)–(15c) and the ranking in (15d).

(15) a. MaxFoot: every foot in the input corresponds to a foot in the output.
b. Dep: Every element in the output has a correspondent in the input (no epenthesis).
c. Integrity: No morph in the input has multiple correspondents in the output.
d. MaxFoot >> Dep >> Integrity (Int)

The ranking of the constraint Dep (banning epenthetic material) over Integrity (banning the re-use of morphs) ensures that underspecified morphemes surface as copies of the base. The trochaic template of reduplication is the result of constraints on foot structure that are active elsewhere in the grammar of German, e.g. for the expression of plural (Eisenberg 1991; Wegener 2004; Wiese 2009), hypocoristics (Féry 1997; Itô and Mester 1997; Wiese 2001), and diminutives (Fanselow and Féry 2002). The following Tableau 1 depicts the emergence of reduplicative morphology in such a case (for ease of exposition, the factors determining the prosodic size of the reduplicant and leading to segmental alternation between base and reduplicant are ignored in this Tableau; see below for more details).

Non-exponence of the expressive morpheme (candidate a.) is infelicitous as it induces a fatal violation of MaxFoot. The affixation of epenthetic material *bla* (candidate b.) is not an option due to Dep. However, the violation of low-ranking Integrity is acceptable, as witnessed by the grammaticality of the

Tableau 1: The grammar of reduplication.

/mɪʃ/$_{\Sigma}$ + /Σ/			MaxFoot	Dep	Int
a.		mɪʃ	!*		
b.		mɪʃ-bla		!*	
c.	☞	mɪʃmaʃ			*

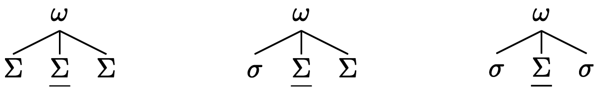
reduplicated *Mischmasch* [mɪʃmaʃ].[19] Further conditions on reduplication in German will be discussed in turn.

4.1 Prosodic shape constraints

The most important constraints on the output in the case of rhyme and ablaut reduplication are the following. First, the output prosodic word is strictly bipedal, and no unfooted syllables are allowed; non-trochaic polysyllabic bases cannot reduplicate. This state of affairs suggests that MaxFoot – and by transitivity Dep and Integrity as well – are dominated by constraints that regulate the prosodic structure of the reduplication. For this purpose, I adopt here familiar constraints on foot structure (Footform = Trochee, requiring feet to be trochaic, i.e. surfacing as bimoraic monosyllable or a disyllable with strong-weak pattern), parsing of syllables (Parse-σ, militating against unfooted syllables), and the notion of Hierarchical Alignment (discussed in Itô et al. [1996] and elaborated on in Ussishkin [2000]). Hierarchical Alignment requires each prosodic constituent to be edge-aligned with some prosodic constituent containing it. In the examples in (16), the edges of prosodic words (ω) to which the feet are properly aligned are marked with "‖" to the left or right, respectively. These structures thus obey Hierarchical Alignment. In contrast, (17) lists several structures, in which the underlined feet fail to align to an edge of a prosodic word, incurring a violation of Hierarchical Alignment.

(16)


19 I assume the segmental string *bla* in candidate b. to be epenthetic material with no phonological relation to material of the stem. In contrast, *masch* in candidate c. holds a specific phonological relation to the segmental structure of the stem and is thus not deemed epenthetic.

(17) 

Alignment to the edge (left or right) of some larger constituent amounts to the expression of prominence within that larger constituent, and, more importantly, obeying this requirement gives rise to maximally binarily branching prosodic words – without the stipulation of a binary maximum. Specifically, Hierarchical Alignment ensures that polypedal bases cannot reduplicate even if the involved feet do not violate the trochaic requirement: in (18), the failure of the innermost feet (underlined) to align to an edge of the prosodic word leads to non-exponence of the expressive reduplicant in the case of *Kunigunde*.

(18) **Kunigundepunigunde* < *Kunigunde* (proper name)
*[[kuː.ni]$_\Sigma$[gʊn.də]$_\Sigma$[puː.ni]$_\Sigma$[gʊn.də]$_\Sigma$]$_\omega$ < [[kuː.ni]$_\Sigma$[gʊn.də]$_\Sigma$]$_\omega$.

For ease of exposition, we assume a prosodic meta-constraint Prosody dominating MaxFoot. This meta-constraint entails the above constraints, namely Parse-σ, Trochee and Hierarchical Alignment. For the present case of reduplication, nothing hinges on their exact intrinsic ordering.

4.2 Segmental non-identity

Apart from the strict prosodic requirements, it was shown that the two feet must not be fully identical. Non-identity is achieved by rhyme or ablaut, depending on the segmental makeup of the base. Bases with more than one full vowel or vowels other than /i/, /ɪ/, /a/, /o/, /ɔ/ can only be reduplicated via rhyme reduplication. However, which technique is used is not fully predictable, as some bases allow both rhyme or ablaut reduplication (*Sillepille*, *Sillesalle* < *Sille*). To cover the nonidentity requirement, I will make use of two constraints (19a) and (19b) proposed for the analysis of haplology in German and English (Plag 1998), and extend their use to the foot level. These two constraints are context-sensitive versions of the more general markedness constraint (19c). Together with (20), these constraints guarantee that adjacent feet be distinct in either the onset or the nucleus of the head syllable.

(19) a. Ocp$_{nucleus}$: the nuclei of (the head syllables of) adjacent feet must not be identical.
b. Ocp$_{onset}$: the onsets of (the head syllables of) adjacent feet must not be identical.
c. Ocp: adjacent feet must not be identical.

Tableau 2: The grammar of rhyme and ablaut reduplication.

			Prsdy	Max$_{\Sigma}$	Id$_{F}$	Ocp	Ocp$_{o}$	Ocp$_{n}$
/mɪʃ/$_{\Sigma}$ + /Σ/								
a.		[mɪ[mɪʃ]$_{\Sigma}$]$_{\omega}$	*!	*				
b.		[[mɪʃ]$_{\Sigma}$]$_{\omega}$		*!				
c.		[[mɪʃ]$_{\Sigma}$][blaː]$_{\Sigma}$]$_{\omega}$			**!*			
d.		[[mɪʃ]$_{\Sigma}$[mɪʃ]$_{\Sigma}$]$_{\omega}$				*	*	*!
e.	☞	[[mɪʃ]$_{\Sigma}$[maʃ]$_{\Sigma}$]$_{\omega}$			*		*	
f.		[[mɪʃ]$_{\Sigma}$[paʃ]$_{\Sigma}$]$_{\omega}$			**!			
/ʃɪki/$_{\Sigma}$ + /Σ/								
a.		[ʃɪ[ʃɪki]$_{\Sigma}$]$_{\omega}$	*!	*				
b.		[[ʃɪki]$_{\Sigma}$]$_{\omega}$		*!				
c.		[[ʃɪki]$_{\Sigma}$][blaː]$_{\Sigma}$]$_{\omega}$			**!*			
d.		[[ʃɪki]$_{\Sigma}$[ʃɪki]$_{\Sigma}$]$_{\omega}$				*	*	*!
e.	☞	[[ʃɪki]$_{\Sigma}$[mɪki]$_{\Sigma}$]$_{\omega}$			*			*
f.		[[ʃɪki]$_{\Sigma}$[maki]$_{\Sigma}$]$_{\omega}$			**!			

(20) Ident-IO$_{F}$: The segmental features of the output are identical to the ones specified in the input.

Tableau 2 depicts the workings of Prosody, MaxFoot, and the interaction of the three Ocp constraints with featural faithfulness Ident-IO$_{F}$ (lower ranking Dep and Integrity are ignored here for ease of exposition); this grammar derives both ablaut reduplication (upper half, *Mischmasch* < *misch*) and rhyme reduplication (lower half, *Schickimicki* < *schick(i)*). The constraint hierarchy prevents reduplication resulting in words with unparsed syllables (candidates a.) or non-exponence of the expressive foot (candidates b.). Filling the underspecified foot with epenthetic material is ruled out (candidates c.), as is total, i.e. identical reduplication (candidates d.). Tying Ocp and Ident-IO$_{F}$, and ranking them above the likewise tied Ocp$_{onset}$ and Ocp$_{nucleus}$, prevents simultaneous application of rhyme and ablaut, cf. the opaque candidates f. The juxtaposition of rhyme and ablaut reduplication in this Tableau demonstrates that both types of reduplication are subject to the same basic constraint hierarchy.

Note that the Ocp constraints are generally violated in ICC and recursive compounds (see Section 2.2.1), while rhyme and ablaut reduplication strictly observe the non-identity requirement. The reason for this discrepancy lies in the nature of the input. Take, for example, a recursive compound like *Kindeskind*, lit.: ‘child of the child’, ‘grandson’; given the compositional transparency of this compound, I assume the input */Kind/ + /Kind/*. A high-ranking constraint MAX-IO (requiring that the output entails the complete featural specification of the input) will prohibit OCP to make any change to the onset or stem vowel (**Kindespind*, **Kindeskand*). In contrast, in the case of the unspecified expressive morpheme in rhyme and ablaut reduplication, MAX-IO has no say about the realization of the reduplicant, so OCP will decide about its surface appearance.[20]

The constraint ranking in the above Tableaux suggests that rhyme and ablaut reduplication are both grammatical options. In Tableau 2, the unattested rhyme reduplication *Mischpisch* for the base *misch* would fare as well as the grammatical candidate *Mischmasch*; likewise, unattested *Schickischacki* would be considered optimal alongside *Schickimicki*. This result is desirable because there are evidently bases that show both options of reduplication (e.g. *Sillesalle*, *Sillepille* < *Sille*). The choice between rhyme and ablaut reduplication is therefore considered to be made outside the phonology proper, unless there are phonological reasons that prohibit one of these options.[21] One such case is considered in the following subsection.

4.2.1 Dorsal fricative alternation blocks ablaut reduplication

The constraint hierarchy in Tableau 2 entails the imperative of a minimal difference between base and reduplicant (cf. the ungrammaticality of candidates d. and f.). As discussed above, ablaut reduplication is blocked when the ablaut would trigger further segmental alternation in the reduplicant (**Michimachi* [mıçimaxi] < *Michi* [mıçi]). The relevant alternation is generally conceived as an assimilation process in which [+back] vowels trigger the backness of a following, tautomorphemic dorsal fricative (/a/ is considered [+back] in German), see (21).

(21) AGREE$_{[back]}$: The dorsal fricative agrees with a preceding tautomorphemic vowel concerning the feature [+/−back].

As shown in Tableau 3, applying ablaut while observing high ranking Agree$_{[back]}$ necessarily leads to a fatal violation of IDENT-IO$_F$ due to the alterna-

20 Assuming, as expounded above, that they are a special kind of endocentric compound, ICCs likewise do not adhere to the prosodic constraints that are active in rhyme and ablaut reduplication; generally, in compounds, the compound members are segmentally specified in the input and given high ranking MAX-IO constraints, they are immune to prosodic shape constraints.

21 This kind of indeterminacy is certainly not unique to reduplication, witness the coexistence of diminuntive affixes -*chen* and -*lein* in German (*Röslein*, *Röschen* 'rose-dim' < *Rose*).

Tableau 3: Dorsal fricative alternation blocks ablaut reduplication.

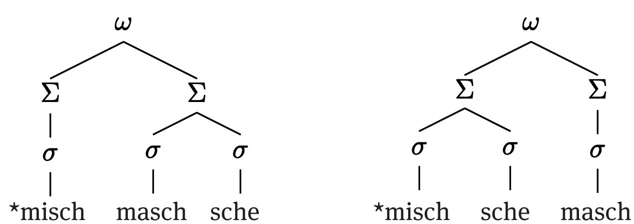
/mıçi/$_{\Sigma}$ + /Σ/			AGREE$_{[back]}$	MAX$_{\Sigma}$	ID$_{F}$	OCP	OCP$_{o}$	OCP$_{n}$
a.	☞	[[mıçi]$_{\Sigma}$[pıçi]$_{\Sigma}$]$_{\omega}$			*			*
b.		[[mıçi]$_{\Sigma}$[mıçi]$_{\Sigma}$]$_{\omega}$				*	*	*!
c.		[[mıçi]$_{\Sigma}$[maxi]$_{\Sigma}$]$_{\omega}$			**!		*	
d.		[[mıçi]$_{\Sigma}$]$_{\omega}$		*!				
e.		[[mıçi]$_{\Sigma}$[maçi]$_{\Sigma}$]$_{\omega}$	*!		*		*	
f.		[[mıxi]$_{\Sigma}$[maxi]$_{\Sigma}$]$_{\omega}$	*!		**		*	

tion of the dorsal fricative (candidate c). Correspondingly, the winning candidate a. is a rhyme reduplication.

4.3 Prosodic parallelism

A conceivable alternative to segmental variation (rhyme or ablaut) as guarantor of non-identity is excluded in reduplication: bipedal forms like (22) with one branching and one non-branching foot are illegitimate reduplications. That is, a difference between base and reduplicant concerning the prosodic shape is prohibited.

(22)

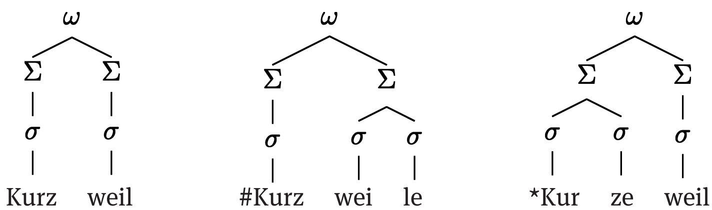


The facts in (22) can be related to a constraint on prosodic parallelism (23) that was recently proposed by Wiese and Speyer (2015). PROSODIC PARALLELISM calls for a symmetric makeup of adjacent prosodic subconstituents (e.g. prosodic words within a prosodic phrase).

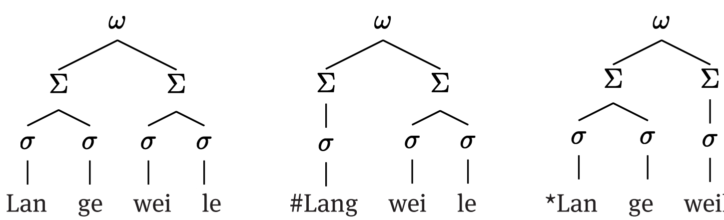
(23) PROSODIC PARALLELISM: Adjacent prosodic subconstituents are symmetric.

Evidence for this constraint comes from various instances of schwa-alternation in German morphophonology. A case in point is the contrast in (24) featuring compounds including the head noun *Weil(e)* (engl.: 'while').[22] In this case, as depicted, the schwa-appearance on the head of the respective compound correlates with schwa-appearance on the modifier. The resulting compound thus obeys prosodic parallelism. Numerous examples of this sort are listed in Wiese and Speyer (2015).

(24) a. *Kurzweil* 'pastime', lit.: 'short while'



b. *Langeweile* 'boredom', lit.: 'long while'



A constraint PROSODIC PARALLELISM is obviously suitable to prohibit ungrammatical reduplications like (22) and, as it has been shown to be effective beyond reduplication, it seems specifically preferable within an account that renounces reduplication-specific constraints that would explicitly call for base-reduplicant correspondence (cf. arguments in Saba Kirchner 2010). As the exact rank of PROSODIC PARALLELISM in the constraint hierarchy cannot be determined here, we may subsume it under the meta-constraint PROSODY.

4.4 Ablaut and morph order

The order base before reduplicant is observed in all rhyme reduplications but not necessarily in ablaut reduplication. To capture this fact, I use the ABLAUT constraint (25) ensuring the order of [+high, front] before [–high, back] among the stem vowels.

22 I thank Birgit Nutz, to whom I owe these examples.

Tableau 4: The grammar of prefixing ablaut reduplication.

/tsakə/$_\Sigma$ + /Σ/	MAX$_\Sigma$	OCP	ABLAUT	LINORDER	INT
a. ☞ [[tsɪkə]$_\Sigma$[tsakə]$_\Sigma$]$_\omega$				*	*
b. [[tsakə]$_\Sigma$[tsɪkə]$_\Sigma$]$_\omega$			*!		*
c. [[tsakə]$_\Sigma$[tsakə]$_\Sigma$]$_\omega$		*!			*
d. [[tsakə]$_\Sigma$]$_\omega$	*!				

(25) ABLAUT: If two feet, one with the high front stem vowel (/i/ or /ɪ/) and one with the non-high back stem vowel (/a/, /o/, or /ɔ/), are conjoined within a prosodic word, the foot with the high front vowel precedes the one with the non-high back vowel.

The ABLAUT constraint interacts with a stipulated constraint LINEAR ORDER. This constraint demands that the exponent most faithful to the stem precedes the exponent of the unspecified morpheme.

(26) LINEAR ORDER: Prosodic constituents that are segmentally specified precede segmentally unspecified prosodic constituents.

That is, if the base features /a/ as stem vowel, the reduplicant with /ɪ/ will be prefixed (*zick(e)zack(e)* < *Zacke*), incurring an acceptable violation of LINEAR ORDER while, in the case of bases with /ɪ/, the reduplicant will be suffixed. Note that the violation of ABLAUT is only crucial in the case of underspecified input. High-ranking faithfulness constraints concerning the linearization of the segmental input (LINEARITY, McCarty and Prince 1995) will spare bipedal stems like *harakiri* from reordering to **hirikara* or **kirihara*. As LINEARITY has no say regarding segmentally empty morphemes, ABLAUT may decide about the order of feet in these cases.

As for rhyme reduplications, the order of faithful and deviant exponent is fixed by LINEAR ORDER, since ABLAUT is irrelevant in the case of two feet with identical stem vowels.

4.5 Non-exponence of the expressive morpheme

The findings of the internet search and the questionnaire attest that non-trochaic bases cannot reduplicate (with the possible exception of dactylic bases, witness ?*Nataliepatalie*). That is, the underspecified foot corresponding to the expressive morpheme does not surface when the base fails to satisfy PROSODY, hence

Tableau 5: Non-exponence of expressive morpheme in the case of non-trochaic bases.

/i'vɔn/ + /Σ/			IdStress	Prosody	Max$_{Σ}$	Ocp
a.		[ʔi[ˈvɔn]$_{Σ}$ pi[ˈvɔn]$_{Σ}$]$_{ω}$		**!$_{Parse-σ}$		
b.		[[ʔiˈvɔn]$_{Σ}$ [piˈvɔn]$_{Σ}$]$_{ω}$		**!$_{Trochee}$		
c.		[[ˈʔivɔn]$_{Σ}$ [ˈpivɔn]$_{Σ}$]$_{ω}$	*!			
d.		[ʔi[ˈvɔn]$_{Σ}$ʔi[ˈvɔn]$_{Σ}$]$_{ω}$		**!$_{Parse-σ}$		*
e.		[ʔi[ˈvɔn]$_{Σ}$ [ˈpivɔn]$_{Σ}$]$_{ω}$		*$_{Parse-σ}$ *!$_{Parallel}$		
f.	☞	[ʔi[ˈvɔn]$_{Σ}$]$_{ω}$		*$_{Parse-σ}$	*	

the ungrammaticality of **Ivonnepivonne* < *Ivonne*. The model predicts that any prosodic deviance, be it unparsed syllables or non-native foot structure, leads to ineffability of rhyme/ablaut reduplication. This is because reduplication will invariably repeat and thus double the prosodic badness of the form (cf. candidates a., b., d. in Tableau 5). As the ungrammaticality of the fully trochaified **'Ivonne'pivonne* (cf. candidate c. in Tableau 5) shows, it is impossible to alter the lexical prosodic specification of the stem in order to make it trochaic and thus reduplicable. Presumably, a high-ranking faithfulness constraint (Ident Stress) prevents the deviance from word stress specified in the input. If the reduplicant abides by the trochaic ideal while the base keeps its lexically specified iambic shape to avoid a violation of Ident Stress,[23] Parallelism and Parse-*σ* will conspire to prohibit the candidate from being realized (cf. candidate e). The winning candidate is therefore the one showing non-exponence of the expressive morpheme.

5 Discussion

The previous section provides an OT approach to the grammar of German rhyme and ablaut reduplication. The analysis does not treat reduplication as a morphological operation but as a phenomenon emerging within the phonology. Accordingly, reduplication surfaces in the face of segmentally underspecified morphs, and it does so without assuming a specific reduplicative morpheme or morphological process calling for segmental copying.

Summarizing the most important points, MaxFoot is dominated by a bundle of constraints (Prosody for short) which govern the prosodic makeup of

23 Strictly speaking, the present purely phonological model cannot distinguish between base and reduplicant as it is blind to morphological affiliation.

the reduplication. This ranking prohibits reduplication of non-trochaic stems and leads to non-exponence of the expressive morpheme in these cases (as evidenced by the ungrammaticality of **Ivonnepivonne*). The ranking in (27a) leads to strictly bipedal, trochaic reduplications, ensuring the non-identity of base and reduplicant. (27b) covers the variable ordering of faithful and deviant exponent in ablaut reduplication while ensuring fixed order in rhyme reduplications. Finally, the ranking in (27c) derives the non-viability of ablaut reduplication when this would lead to additional segmental alternations (**Michimachi*). It also reflects the ban of opaque reduplications that apply both rhyme and ablaut at the same time.

(27) a. PROSODY >> MAX_{Σ} >> OCP >> DEP >> INTEGRITY
b. MAX_{Σ} >> ABLAUT >> LINEAR ORDER
c. AGREE_{back} >> MAX_{Σ} >> IDENT-IO_{F}, OCP >> OCP_{onset}, $\text{OCP}_{nucleus}$

This OT grammar was shown to successfully model the emergence of the fixed bipedal structure, the obligatory segmental deviance of the reduplicant, non-exponence of the expressive morpheme in the case of non-trochaic bases, the variable linearization of base and reduplicant in ablaut reduplication, and the interaction of reduplication with segmental alternations.

Several aspects of reduplication, however, remain unaccounted for. For one thing, it is as yet unclear what exactly conditions the choice between ablaut and rhyme reduplication. The evaluation of the collection of rhyme and ablaut reduplication provides some hints: the avoidance of complex onsets and the bias towards disyllabic bases in the case of rhyme reduplication shows that the segmental and syllabic makeup of the base co-determines the kind of reduplication. However, cases of optionality (*Sillesalle* ~ *Sillepille* < *Sille*) show that the grammar does not enforce the choice between these structures.

Secondly, the present OT model, while predicting the impossibility of reduplication in the case of non-trochaic bases, does not offer a solution as to how the grammar might resort to other ways to reveal the semantics associated with the expressive morpheme. Note that the findings of the questionnaire suggest that reduplication is fairly acceptable with stems that were truncated to a trochee *Susepuse*, *Sannepanne*, $^{?/}$**Susánnepusànne* < *Susanne*. Forms like *Susepuse* < *Susanne* may be considered “templatic backcopies”, in which the base is truncated to match the prosodic form of the truncated reduplicant.[24] Truncation

24 Interestingly, templatic backcopying in reduplication has been argued to be nonexistent in natural languages (Spaelti 1999, among others). The German data presented here, along with data from Guarijio (Caballero 2006), clearly attest to the existence, and thus refute the alleged impossibility of backcopying.

is a well established process of German prosodic morphology. The most common way of truncation is hypocoristic -*i*-formation (Féry 1997; Grüter 2003; Itô and Mester 1997; Wiese 2001). In contrast to reduplication, for *i*-truncations, the prosody of non-trochaic stems may be tweaked to fit a trochaic template. A case in point is the nickname *Ivi* ['ʔivi] < *Ivonne*. Interestingly, the output of *i*-truncation may well become the input of reduplication, as witnessed by *Ivipivi* < *Ivi* < *Ivonne*. That is, the faithfulness constraints (IDENTSTRESS) that dominate constraints on prosodic structure in the case of reduplication are themselves dominated by prosodic constraints in the case of -*i*-truncation. This state of affairs strongly suggests different constraint rankings for the two morphological processes. A possible solution to this dilemma would be a stratal organisation of grammar with different cyclic domains for *i*-truncation and reduplication (Bermúdez-Otero 2012). Alternatively, either one or the other process has to be deemed outside grammar proper – Bye and Svenonius (2012) opt for the latter and suggest truncation to be extragrammatical in the sense of Dressler (2000). In any case, with the paradoxical ranking of IDENTSTRESS and PROSODY, it seems to be impossible to model both reduplication and templatic truncation within a single grammatical stratum.

Before concluding, I note in the following paragraph certain cross-linguistic correspondences between reduplication and the grammar of metered verse in German(ic) and Romance.

5.1 Reduplication reflects versification and poetic rhyme

The fact that reduplication is mainly used for expressive purposes in playful registers and the term "rhyme reduplication" are suggestive of the poetic dimension of language. It may not come as a surprise then, that we find certain relations between reduplicative word formation and poetry. In this context, it is interesting to note that, while rhyme and ablaut reduplication in German(ic) are based on the prosodic foot, resulting in either disyllabic or quadrisyllabic reduplications, reduplication in French is strictly syllable-based, resulting in disyllabic words (e.g. *dodo* < *dormir* 'to sleep', *gaga* < *gateux* 'crazy') (Scullen 2002; Lambert 2004). Likewise, Italian reduplications, which may be found in reduplicating truncations (Alber 2010), always result in disyllables (*Gigi* < *Luigi*, *Lele* < *Elena*). This crosslinguistic difference is reflected in the versification of the respective languages: While German poetry counts stresses i.e. feet, leaving some freedom regarding the number of syllables in a verse, French and Italian poetry are more strictly based on the number of syllables. Any deviance from the syllable number determined by the poetic meter is considered infelicitous in these languages.

The cross-linguistic correspondence concerning the prosody of poetry and reduplication is enhanced by a segmental correspondence between these domains. That is, French and Italian reduplications often preserve the identity of base and reduplicant and, at the same time, these languages allow identical rhymes in poetry (e.g. Aroui 2005). In contrast, German(ic) generally disallows identical reduplication and, coincidentally, identical rhymes in poetry are clearly considered unsatisfactory. A correspondence of this kind is most probably not accidental.

However, the observation of the prosodic and segmental correspondence in word formation and poetry does not explain why there is identity avoidance between corresponding feet (Germanic) and identity observance between corresponding syllables (French, Italian). I will leave this issue open at this point. Note however that, to the marginal extent that syllabic reduplication is active in German (cf. the cases of phonological doubling in (4) and a couple of nicknames: *Vivi* < *Viola*, *Kiki* < *Kirsten*), the identity of the syllables is observed, too.

6 Conclusion

I have provided evidence for the claim that reduplication in German deserves a systematic treatment and is amenable to a formal analysis. The perspective on reduplicative word formation offered here provides new insights into German morphophonology. The proposed taxonomy of this diverse phenomenon takes its cue from the lexicality of the form and leads to a delineation which identifies rhyme and ablaut reduplication as the only productive, truly reduplicative processes in the morphology of German. All other contemplable cases are properly treated as either phonological doubling, lexical sequencing or (special cases of) compounding.

The informal corpus search and the acceptability rating study attest the productivity and regularity of rhyme and ablaut reduplication and thus call for a formal treatment of these processes. The OT analysis demonstrates that both rhyme and ablaut reduplication may emerge when a segmentally and prosodically underspecified expressive morpheme is attached to a base – given that the base strictly obeys certain prosodic requirements, esp. concerning its foot structure. The present approach to reduplication eschews constraints that make explicit reference to base-reduplicant correspondence. Rather, reduplication is shown to be a special case of concatenative morphology, where the morpheme attached to the base is a plagiarist that avails itself of the segmental material of the base, making only slight changes to the original.

The OT grammar submitted here successfully models the emergence of the fixed bipedal structure, the obligatory but minimal segmental deviance of the reduplicant, the non-exponence of the expressive morpheme in the case of non-trochaic bases, the variable ordering of base and reduplicant in ablaut reduplication, and the interaction of reduplication with segmental alternations (blocking of ablaut reduplication in the face of [ç-x]-alternation).

I leave for future work the interaction of i-truncation and reduplication that poses several problems. Wiese (2001) has emphasized the kinship of these two instances of prosodic morphology. However, the kinship is a complicated one: How is it that the same grammar that blocks reduplications of non-trochaic bases (**Ivonnepivonne* < *Ivonne*) is able to trochaify exactly those bases to build *i*-truncations (*Ivi* < *Ivonne*), which can then serve as the input to reduplication (*Ivipivi* < *Ivi*)?

To conclude, while reduplication may be considered marginal in German (because German linguistics has assigned it a marginal status), it is certainly not to be taken as (typologically) exceptional or rare (cf. Dingemanse [2017] for a distinction between marginalia and rara; Joseph [1997] on the importance of marginalia). Instead, I hope to have shown that reduplication systematically adheres to grammatical constraints that are known to be active in the morphophonology of German (and elsewhere), and is thus part of German speakers' linguistic competence.

Acknowledgments: This work has benefitted from discussions at several occasions, most notably the P&P conference in Osnabrück (2011) and the Phonology 2013 conference in Amherst. Thanks are due to at least three anonymous reviewers whose comments helped improving the paper considerably. Especially, I would like to thank the following friends and colleagues for discussions of (aspects of) this work: Beata Moskal, Caroline Féry, Eva Zimmermann, Marc van Oostendorp, Markus Hiller, Richard Wiese, Ruben van de Vijver, and Sabine Zerbian. Alex Pösch was immensely helpful with sorting out last-minute formatting issues.

References

Alber, Birgit. 2010. An exploration of truncation in Italian. In Peter Staroverov, Daniel Altshuler, Aaron Braver, Carlos A. Fasola & Sarah Murray (eds.), *Rutgers Working Papers in Linguistics*, Vol. 3, 1–30. New Brunswick, NJ: LGSA, Rutgers University.

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

Aroui, Jean-Louis. 2005. Rime et richesse des rimes en versification fraņcaise classique. In Michel Murat & Jacqueline Dangel (eds.), *Poétique de la rime*, 179–218. Paris: H. Champion.

Barz, Irmhild. 2015. German. In Peter Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe*, Vol. 4: *Word formation in the individual European languages*, 2387–2410. Berlin & Boston: Walter de Gruyter.

Bauke, Leah S. 2014. *Symmetry breaking in syntax and the lexicon*. Amsterdam: John Benjamins.

Benczes, Réka. 2012. Just a load of Hibber-Gibber? Making sense of English rhyming compounds. *Australian Journal of Linguistics* 32. 299–326.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, 8–83. Oxford: Oxford University Press.

Burzio, Luigi. 1994. *Principles of English stress*. Cambridge University Press.

Bye, Patrik & Peter Svenonius. 2012. Non-concatenative morphology as epiphenomenon. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, 427–495. Oxford: Oxford University Press.

Bzdęga, Andrzej. 1965. *Reduplizierte Wortbildung im Deutschen*. Poznań: Poznańskie Tow. przyjaciól nauk.

Caballero, Gabriela. 2006. “Templatic backcopying” in Guarijio abbreviated reduplication. *Morphology* 16. 273–289.

Cooper, William E. & John Robert Ross. 1975. World order. In Robin Grossman, James San & Timothy Vance (eds.), *Papers from the Parasession on Functionalism*, 63–111. Chicago, IL: Chicago Linguistic Society.

Dingemanse, Mark. 2017. On the margins of language: Ideophones, interjections and dependencies in linguistic theory. In Nick Enfield (ed.), *Dependencies in language*, 195–203. Berlin: Language Science Press.

Dressler, Wolfgang U. 2000. Extragrammatical vs. marginal phonology. In Ursula Doleschal & Anna M. Thornton (eds.), *Extragrammatical and marginal phonology*, 2–10. Munich: Lincom.

Eisenberg, Peter. 1991. Syllabische Struktur und Wortakzent. Prinzipien der Prosodik deutscher Wörter. *Zeitschrift für Sprachwissenschaft* 10. 37–64.

Fanselow, Gisbert & Caroline Féry. 2002. Ineffability in grammar. *Linguistische Berichte* Sonderheft 11. 265–307.

Féry, Caroline. 1997. Uni und Studis: die besten Wörter des Deutschen. *Linguistische Berichte* 172. 461–489.

Fikkert, Paula, Marieke van Heugten, Philo Offermans & Tania S. Zamuner. 2005. Rhymes as a window into grammar. In *Proceedings of the 29th Annual Boston University Conference on Language Development*, Vol. 1, 204–215. Sommerville, MA: Cascadilla Press.

Finkbeiner, Rita. 2012. Naja, normal und normal. Zur Syntax, Semantik und Pragmatik der x-und-x-Konstruktion im Deutschen. *Zeitschrift für Sprachwissenschaft* 31. 1–42.

Finkbeiner, Rita. 2014. Identical constituent compounds in German. *Word Structure* 7. 182–213.

Freywald, Ulrike. 2015. Total Reduplication as a productive process in German. *Studies in Language* 39. 905–945.

Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen & Kevin Russell. 2004. Contrastive focus reduplication in English (The salad-salad paper). *Natural Language & Linguistic Theory* 22. 307–357.

Grüter, Theres. 2003. Hypocoristics: The case of u-formation in Bernese Swiss German. *Journal of Germanic Linguistics* 15. 27–63.

Harley, Heidi. 2012. Compounding in Distributed Morphology. In Rochelle Lieber & Pavol Stekauer, *The Oxford handbook of compounding*, 129–144. Oxford: Oxford University Press.

Heiberger, Richard M. & Burt Holland. 2015. *Statistical analysis and data display*. New York: Springer. URL http://CRAN.R-project.org/package=HH (accessed 02 March 2017).

Hohenhaus, Peter. 2004. Identical constituent compounding – a corpus-based study. *Folia Linguistica* 38. 297–331.

Inkelas, Sharon. 2008. The dual theory of reduplication. *Linguistics* 46. 351–401.

Inkelas, Sharon & Cheryl Zoll. 2005. *Reduplication: Doubling in morphology*. Cambridge: Cambridge University Press.

Itô, Junko, Yoshihisa Kitagawa & Armin Mester. 1996. Prosodic faithfulness and correspondence: Evidence from a Japanese argot. *Journal of East Asian Linguistics* 5. 217–294.

Itô, Junko & Armin Mester. 1997. Sympathy theory and German truncations. *University of Maryland Working Papers in Linguistics* 5. 117–139.

Joseph, Brian D. 1997. On the linguistics of marginality: The centrality of the periphery. In Cora Singer, Randall Eggert & Gregory Anderson (eds.), *Proceedings from the 33th Annual Meeting of the Chicago Linguistic Society*, Vol. 1: *The Main Session*, 197–213. Chicago, IL: Chicago Linguistic Society.

Lambert, Renée. 2004. Diminutive reduplication in government phonology. *Revista de Estudos da Linguagem* 12. 89–113.

Lappe, Sabine. 2008. *English prosodic morphology*. Heidelberg: Springer.

Leiner, DJ. 2015. *SoSci Survey [Computer Software]*. URL https://www.soscisurvey.de.

Lisker, Leigh. 1974. On 'explaining' vowel duration variation. *Haskins Laboratories Status Report on Speech Research* 37. 225–232.

Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13. 435–482.

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Laura Dickey, Jill Beckman, and Suzanne Urbanczyk (eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers in Linguistics 18), 249–384. Amherst, MA: GSLA Publications.

Meibauer, Jörg. 2007. How marginal are phrasal compounds? Generalized insertion, expressivity, and I/Q-interaction. *Morphology* 17. 233–259.

Minkova, Donka. 2002. Ablaut reduplication in English: The criss-crossing of prosody and verbal art. *English Language and Linguistics* 6. 133–169.

Müller, Gereon. 1997. Beschränkungen für Binominalbildung im Deutschen – ein Beitrag zur Interaktion von Phraseologie und Grammatik. *Zeitschrift für Sprachwissenschaft* 16. 5–51.

Nevins, Andrew & Bert Vaux. 2003. Metalinguistic, shmetalinguistic: The phonology of shm-reduplication. In Jonathon E. Cihlar, David Kaiser, Irene Kimbara & Amy Franklin (eds.), *Proceedings from the 39th Annual Meeting of the Chicago Linguistic Society*, Vol. 1: *The main session*, 702–721. Chicago, IL: Chicago Linguistic Society.

Plag, Ingo. 1998. Morphological haplology in a constraint-based morphophonology. In Richard Wiese & Wolfgang Kehrein (eds.), *Phonology and morphology of the Germanic languages*, 199–215. Tübingen: Niemeyer.

Primus, Beatrice. 2002. Unreine Reime und phonologische Theorie. In David Restle & Dietmar Zaefferer (eds.), *Sounds and systems. Studies in structure and change. A festschrift for Theo Vennemann*, 269–298. Berlin & Boston: de Gruyter.

Roeper, Thomas, William Snyder & Kazuko Hiramatsu. 2002. Learnability in a Minimalist framework: Root compounds, merger, and the syntax-morphology interface. In Ingeborg Lasser (ed.), *The process of language acquisition*, 25–37. Frankfurt: Peter Lang.

Ronneberger-Sibold, Elke. 2006. Lexical blends: Functionally tuning the transparency of complex words. *Folia Linguistica* 40. 155–181.

Rubino, Carl. 2011. Reduplication. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Munich: Max Planck Digital Library. URL http://wals.info/chapter/27.

Saba Kirchner, Jesse. 2010. *Minimal reduplication*. Santa Cruz, CA: University of California dissertation.

Schindler, Wolfgang. 1991. Reduplizierende Wortbildung im Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 44. 597–613.

Scullen, Mary Ellen. 2002. New insights into French reduplication. In Caroline R. Wiltshire & Joaquim Camps (eds.), *Romance phonology and variation*, 177–190. Amsterdam: Benjamins.

Spaelti, Philip. 1999. *Dimensions of variation in multi-pattern reduplication*. Santa Cruz, CA: University of California dissertation.

Stolz, Thomas. 2008. Total reduplication vs. echo-word formation in language contact situations. In Peter Siemund & Noemi Kintana (eds.), *Language contact and contact languages*, 107–132. Amsterdam: Benjamins.

Torre, Erik Jan van der. 2003. *Dutch sonorants. The role of place of articulation in Dutch phonotactics*. Leiden: University of Leiden dissertation.

Ussishkin, Adam Panter. 2000. *The emergence of fixed prosody*. Santa Cruz, CA: University of California dissertation.

Wegener, Heide. 2004. Pizzas und Pizzen – die Pluralformen (un)assimilierter Fremdwörter im Deutschen. *Zeitschrift für Sprachwissenschaft* 23. 47–112.

Wierzbicka, Anna. 2003. *Cross-cultural pragmatics*. Berlin & Boston: de Gruyter.

Wiese, Richard. 1990. Über die Interaktion von Morphologie und Phonologie – Reduplikation im Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 43. 603–624.

Wiese, Richard. 1996. *The phonology of German*. Oxford: Clarendon Press.

Wiese, Richard. 2001. Regular morphology vs. prosodic morphology? The case of truncations in German. *Journal of Germanic Linguistics* 13. 131–177.

Wiese, Richard. 2009. The grammar and typology of plural noun inflection in varieties of German. *The Journal of Comparative Germanic Linguistics* 12. 137–173.
Wiese, Richard & Augustin Speyer. 2015. Prosodic parallelism explaining morphophonological variation in German. *Linguistics* 53. 525–559.

Appendix

List of reduplications

A.1 Ablaut reduplication

A.1.1 Prefixing ablaut reduplication

flickerflacker < flacker
flitterflatter < flatter
Gimmelgammel < Gammel
Hickhack < hack
Kitzekatze < Katze
Krikelkrakel < krakel
Krimskrams < Kram
Mimpfmampf < Mampf
Mipsmops < Mops
Mitschmatsch < Matsch
Mitzematze < Matze
Mitzmatz < Matz
nigelnagel(neu) < Nagel
pickepackevoll < pack
pipelpopel < Popel
pitschpatsch < patsch
plipperplapper < plapper
plitschplatsch < platsch
Quitschquatsch < Quatsch
rickerracker < Racker
Rilleralle < Ralle ~ Ralf
rimmelrammel < rammel
rischelraschel < raschel
ritzirotzi < Rotz
schnickschnack < schnack
schwibbelschwabbel < schwabbel
schwippschwapp < schwapp
Stipstop < stop
tipptopp < top
tiptap < tapp

Zickezacke < Zacke
zippelzappel < zappel
Zittelzottel < Zottel

A.1.2 Suffixing ablaut reduplication

Brillebralle < Brille
fimmelfammel < Fimmel
Flügelflagel < Flügel
giggelgaggel < giggel
Kippelkappel < kippel
Knister-knaster < knister
Krillekralle < Krille ~ Christian
Kritzelkratzel < kritzel
Mischmasch < misch
Mixmax < Mix
Pickelpackel < Pickel
Pimmelpammel < Pimmel
prickelprackel < prickel
Prinzpranz < Prinz
Rikerake < Rike
ritzeratze < Ritze
Schlingelschlangel < Schlingel
Singsang < sing
schwingschwang < schwing
Sillesalle < Sille ~ Silke
Stinkstonk < stink
ticktack < tick
Tingeltangel < tingel
Wiebkewabke < Wiebke
Wirriwarri < wirr
Wirrwarr < wirr
Zipfelzapfel < Zipfel
Zwickzwack < zwick

A.2 Rhyme reduplication

Achimpachim < Achim
Andimandi < Andi
Andipandi < Andi
Annapanna < Anna
Annipanni < Anni
Binemine < Bine ~ Sabine
Binepine < Bine ~ Sabine

Buckelmuckel < Buckel
Compipompi < Compi ~ Computer
Daddelpaddel < daddel
Danimani < Dani ~ Daniel, Daniela
Danipani < Dani ~ Daniel, Daniela
dengelbengel < dengel
Doppelmoppel < doppel
Franzipanzi < Franzi
Fritzefitze < Fritz(e)
Furzipurzi < Furz(i)
Fusselpussel < Fussel
Hasemase < Hase
HaukePauke < Hauke
Heikepeike < Heike
Heinzipeinzi < Heinz(i)
Heinzpeinz < Heinz
Hennapenna < Henna
Hinkepinke < hink
Horstiporsti < Horst(i)
Ilsebilse < Ilse
Inamina < Ina
Inapina < Ina
Ingopingo < Ingo
Kallepalle < Kalle ~ Karsten
Karloparlo < Karlo
kitzelbitzel < kitzel
Klausipausi < Klaus(i)
Kuschelmuschel < kuschel
Kuschelpuschel < kuschel
Larsiparsi < Lars(i)
Mannipanni < Manni ~ Manfred
Manupanu < Manu ~ Manuel, Manuela
Matzepatze < Matze ~ Matthias
Matzpatz < Matz ~ Matthias
mausipausi < Maus(i)
meikepeike < Meike
Michipichi < Michi ~ Michael, Michaela
Miekepieke < Mieke
Mietzepietze < Mietze
mogelpogel < mogel
motzipotzi < motz(i)
müffelpüffel < müffel
Muschipuschi < Muschi
Nickipicki < Nicki ~ Nicole, Nikolas
okidoki < OK
Pepemepe < Pepe
Popelmopel < Popel
Pupsidupsi < Pups(i)

Putzimutzi < putz
Ralfipalfi < Ralf(i)
Ralfmalf < Ralf
Ralfpalf < Ralf
rammelpammel < rammel
Rikepike < Rike
Rotzipotzi < Rotz(i)
Rotzpotz < Rotz
Rubbeldubbel < rubbel
Schickimicki < schick(i)
schlingelpingel < Schlingel
Schorlemorle < Schorle
Schusselpussel < Schussel
Sebimebi < Sebi ~ Sebastian
Silkepilke < Silke
Sillepille < Sille ~ *Silke*
Sinebine < Sine ~ Sina, Sabine
Steffenpeffen < Steffen
Stinkipinki < stink
Stoffelpoffel < Stoffel
Stöpselpöpsel < Stöpsel
superduper < super
Susepuse < Suse ~ Susanne
Susipusi < Susi ~ Susanne
Thilopilo < Thilo
Trippelpippel < trippel
Udopudo < Udo
Ulfipulfi < Ulf(i)
Ullipulli < Ulli ~ Ulrich, Ulrike
Veramera < Vera
Wickeldickel < wickel
Wimmelbimmel < wimmel
Wollebolle < Wolle ~ Wolfgang
Wollemolle < Wolle ~ Wolfgang
Wollepolle < Wolle ~ Wolfgang
wuselpusel < wusel
Zottelmottel < Zottel
Zottelpottel < Zottel
zuckelmuckel < zuckel

Gerrit Kentner

Schwa optionality and the prosodic shape of words and phrases

Abstract: This paper discusses effects of supra-lexical linguistic rhythm on the appearance or absence of optional schwa. Specifically, the roles of rhythmic alternation and prosodic parallelism are studied in three experiments and weighed against each other. In Experiment One, an oral reading study, readers were confronted with either of the two graphemic representations of the alternating adverb <gern(e)> ('happily') in sentential contexts the rhythmic structure of which was systematically varied. The evaluation of the scripted speech productions suggests that readers take the rhythmic environment into account when choosing an allomorph for the prosodically variable target word. Experiment Two is concerned with prosodic determinants for the morphosyntactic alternation in German partitive or possessive constructions. These may be realised as genitive attributes or using a prepositional construction. A forced choice experiment with written material suggests that participants consider the distribution of strong and weak syllables when choosing among the morphosyntactic variants. Experiment Three exploits the prosodic alternation of four adverbs. Analysing the distribution of the variants in a large written corpus attests that the immediate prosodic context affects the choice among the variants. A synopsis of the findings suggests that rhythmic alternation (conceived as the joint effects of stress clash avoidance and stress lapse avoidance) has a stronger impact on the presence or absence of a reduced syllable compared to prosodic parallelism.

Keywords: word prosody, optional schwa, prosodic morphology, German

Gerrit Kentner: Johann Wolfgang Goethe-Universität, Norbert-Wollheim-Platz 1, D—60629 Frankfurt am Main, kentner@lingua.uni-frankfurt.de

DOI 10.1515/9783110542899-006

1 Introduction

Several words in German are prosodically variable in that they may legally appear either with or without a schwa syllable – with no semantic effect associated with the presence or absence of the syllable headed by the reduced vowel. Schwa optionality is certainly a marginal phenomenon in German morphophonology, most likely because the appearance of schwa is usually morphologically governed, with schwa corresponding to the exponent of e.g. a plural, first person, or agreement morpheme. In spite of schwa's role as inflectional exponent, schwa optionality is attested in all major word classes, as the list in (1) reveals.

(1) a. Nouns: *die Tür ~ die Türe* ('the door')
b. Verbs (first person sg., pres.): *ich geh ~ ich gehe* ('I go')
c. Predicative adjectives: *fad ~ fade* ('dull, tasteless')
d. Adverbs: *gern ~ gerne* ('happily')
e. Demonstrative pronouns: *dies ~ dieses* ('this')
f. Conjunctions: *eh' ~ ehe* ('before')
g. Preposition: *ohn' ~ ohne* ('without')
h. Numerals: *zu zweit ~ zu zweien* ('two by two', 'in sets of two')

A variety of factors are known to impinge on the appearance or absence of optional schwa syllables, among them language change, speaking rate and style (or register), and dialectal influence. That is, not all alternating forms in (1) are equally likely to occur in modern Standard German. For instance, the dated numeral (1h.) of the form *zu NUM-en* has by now been almost fully replaced by the current schwa-less construction *zu NUM-t*. Monosyllabic *ohn'* (1g.) is confined to certain poetic registers, while (1c.) has a dialectal distribution. The appearance of optional schwa (or its orthographic cognate <e>) in genitives *Jahrs ~ Jahres* ('year') has been shown to be more likely the higher the frequency of the noun is (Fehringer 2011). Still, some alternating forms seem to happily coexist and vary almost freely even within the same historical and dialectal strata.

Aside from factors like usage frequency, speech register, and dialectal distribution, the prosodic-phonological context the variable word is embedded in has been discussed as potentially conditioning the distribution of forms with or without schwa syllable. Studies by Rohdenburg (2014), Schlüter (2005), and Wiese and Speyer (2015) suggest that the prosodic makeup of adjacent words may co-determine the choice among the prosodically varying allomorphs. The claim put forward in these studies is that speakers exploit schwa-optionality to improve the phrasal rhythm. There are at least two, partly conflicting, ways in which phrasal rhythm may be improved. On the one hand, speakers may, whenever possible, strive for an alternation of stressed (or strong) and unstressed (or weak) syllables, thereby creating a beat that is as regular as possible. This entails that structures involving sequences of adjacent stressed syllables (stress clash) or sequences of unstressed syllables (stress lapses) are disfavoured. On the other hand, the rhythmicity of an utterance may be enhanced by iterating prosodic units of the same type, fostering prosodic parallelism. Accordingly, a prosodic phrase that consists of two words is favoured if the two words exhibit the same prosodic structures (e.g. either two monosyllabic words or two trochees); a sequence of two prosodically different words (e.g. a trochee followed by a monosyllable) would violate the iterative rhythm.

This paper takes a fresh look at the various effects of supra-lexical linguistic rhythm on the appearance or absence of optional schwa. Specifically, the roles of rhythmic alternation on the one hand, and iterative rhythm or prosodic parallelism on the other will be studied in three experiments and weighed against each other. In Experiment One (Section 2.1), an oral reading study, readers were confronted with either of the two graphemic representations of the alternating adverb <gern(e)> ('happily') in sentential contexts that were systematically varied with respect to rhythmic structure. The evaluation of the scripted speech productions suggests that readers take the rhythmic environment into account when reading out the written target word. Experiment Two (Section 2.2) is concerned with prosodic determinants for the morphosyntactic alternation in German partitive or possessive constructions. These may be realised as genitive attributes or using a prepositional construction. A forced choice experiment with written material suggests that participants consider the distribution of strong and weak syllables in the possessum when choosing among the morphosyntactic variants, confirming a rhythmic-prosodic effect. Finally, Experiment Three (Section 2.3) exploits the prosodic alternation of the adverbs *gern* ~ *gerne*, *lang* ~ *lange*, *selbst* ~ *selber*, *meist* ~ *meistens* ('happily, for a long time, self, most of the time'). Analysing the distribution of the variants in a large written corpus attests that the propensity for rhythmic alternation affects the choice among these variants. A synopsis of the endings suggests that rhythmic alternation (conceived as the joint effects of stress clash and stress lapse avoidance) has a stronger effect on the presence or absence of a reduced syllable compared to prosodic parallelism. Before reporting on the experiments in Section 2, the remainder of Section 1 provides relevant background on prosodic structure and linguistic rhythm in German (and beyond).

1.1 (Supra-)lexical prosodic structure and linguistic rhythm

As for word-internal prosody, the core of the German lexicon and morphological system is prosodically constrained in that it displays a strong preference for disyllabic, trochaic forms (for a review, see Domahs, Domahs, and Kauschke 2017, this volume). The trochaic preference dictates e.g. the choice of plural allomorphs (Eisenberg 1991; Wegener 2004; Wiese 2009), and it restricts the productivity of many derivations, such as umlaut in diminutive formation (Fanselow and Féry 2002) or the possibility to form denominal adjectives by suffixation of *-ig*[1] (this derivation is only licit when the suffix is immediately preceded by a syllable carrying stress, thus forming a right-aligned trochee: *ruhig* < *Ruhe*, *tomatig* < *Tomate*, **kürbisig* < *Kürbis*, **mangoig* < *Mango*, **paprikaig* < *Paprika*[2]). The effect of the trochee in German morphology is probably best seen in hypocoristic truncations with the *i*-suffix (*Ándi* < *Andréas*, *Stúdi* < *Studént*) (cf. Féry 1997; Itô and Mester 1997; Köpcke 2002) in which the trochaic template applies almost exceptionless – in fact, as the examples *Andi* and *Studi* show, this highly productive process may even force the deviance from the stress pattern of the source form to safeguard a trochee. The trochee may thus be understood as an optimal template regulating the shape of words.

Beyond the word, the trochee may lead to rhythmic alternation of strong and weak syllables. In the ideal case, the concatenation of words yields a concatenation of trochees and, consequentially, the perfect alternation of strong and weak beats. A trochaic structure like (2) fulfills pertinent conditions regarding rhythmic alternation, namely the constraints against clustering of strong syllables (*CLASH) (see Anttila et al. 2010, for various instantiations of this constraint) or against sequences of weak syllables (*LAPSE) (cf. Shih et al. 2015, for a discussion of different eurhythmy measures). The example in (2) can be considered especially eurhythmical in that the alternation between strong and weak is even reflected in the vowel qualities with diphthongs or long vowels alternating with unstressable reduced vowels.

1 In the case of stems ending in a syllabic sonorant, a schwa syllable may be skipped, producing dactylic forms like *hibbelig* ('jittery'). This is reminiscent of Kager's notion of the invisibility of schwa syllables to certain phonological processes that are conditioned by stress (Kager 1989), as may be exemplified by German umlaut (Féry 1994).

2 A reviewer suggests that hiatus avoidance may be considered a factor in the ungrammaticality of **mangoig* and **paprikaig*. However, cases like *ruhig* [ʁuːɪç] and *böig* [bøːɪç] < *Böe* ('gust', 'squall') attest the license for hiatus in these contexts.

(2) *Friede, Freude, Eierkuchen*
[(ˈfʁiː.də) (ˈfʁɔɪ.də) (ˈʔaɪ.ɐ) (ˌkuː.xən)]
peace, joy, pancake
'love, peace and harmony'

Apart from the alternating rhythm of strong and weak syllables, (2) constitutes a prime example for iterating rhythm (or prosodic parallelism) with the four trochees building a perfectly parallelistic prosodic structure, i.e. a symmetric set of two pairs of trochees. The prosodic repetitiveness is enhanced by the segmental structure at least for the first pair of trochees (the parallelism is mirrored in the onsets of both the stressed and the unstressed syllable). The alliterating idiomatic expression in (2) suggests that the force of iterating rhythm is most obvious in poetic language where prosodic parallelism is prevalent (concerning, for example, the matching of lines in metered poems, see Menninghaus et al. 2017).

1.2 Rhythmic alternation within and beyond the word

The propensity for rhythmic alternation (i.e. the effect of *CLASH and *LAPSE) is illustrated by cases in which it forces a deviance from patterns that would be expected by mere concatenation of morphs. For instance, the prominence of syllables can be demoted to avoid a clash of neighbouring strong beats. Consider, in this respect, the German word *Nation* [naˈtsjoːn] ('nation') with the latinate suffix *-ion* attracting stress on the final syllable. Attaching the equally stress-attracting adjectival suffix *-al* leads to a restructuring of prominences such that the stem-final syllable becomes unstressed and instead the initial syllable receives secondary stress (*national* [ˌnatsjoˈnaːl]).

In other cases, the force of *CLASH may even impinge on the quality of the underlying vowel. This is the case in the most natural rendition of a compound like *Bauarbeiter* 'builder', made up of the constituents *Bau* [baʊ] 'building' and *Arbeiter* [ˈʔaː.baɪ.tɐ] 'worker'. With compound stress on the first constituent, the initial syllable of the second member becomes a reduced syllable and is thus attached to the foot projected by the monosyllabic first member [(ˈbaʊɐ)(ˌbaɪ.tɐ)]. Importantly, the footing of this compound, arguably driven by *CLASH, runs counter to its morphological structure.[3]

3 It is certainly possible for *Bauarbeiter* to retain secondary stress on the first syllable of the head noun and, in addition, mark the morphological boundary by a glottal stop

The avoidance of clashes has also been shown to have syntactic effects (cf. especially Schlüter 2005; Speyer 2010, on the syntactic effects of clash avoidance in English). As for German, consider the otherwise unmotivated ordering with the adverbial intensifier *ganz* or *so* that is separated from the adjective or de-adjectival noun it modifies and instead preceding the indefinite pronoun or determiner in noun phrases like (3) (see e.g. Kallulli and Rothmayr 2008 and Gutzmann and Turgay 2015 for syntactic and semantic analyses of similar phenomena). This inversion coexists alongside the canonical ordering with the determiner preceding the intensifying adverb. As noted by Behaghel (1930), the displaced determiner serves as a buffer between two prominent syllables, preventing a clash. Determiner doubling in (4) provides an even more striking case, arguably with the same motivation.[4]

(3) canonical order ~ determiner inversion
  a. *was ganz Neues ~ ganz was Neues*
     'something quite new' ~ 'quite something new'
  b. *ein ganz junger Mann ~ ganz ein junger Mann*
     'a quite young man' ~ 'quite a young man'

(4) determiner doubling
  a. *ein ganz ein feiner Kerl*
     'a quite a fine chap'
  b. *ein so ein großer Bub*
     'a such a big boy'

Interestingly, inversion or doubling appears to be illicit in German with di- or trisyllabic intensifiers (*gänzlich*, *dermaßen*) whose unstressed final syllable prevents a stress clash in the first place.

(5) a. *ein gänzlich feiner Kerl*
  a.' **gänzlich ein feiner Kerl*
  a." **ein gänzlich ein feiner Kerl*
     '(a) quite (a) fine chap'

[(ˈbaʊ)(ˌʔaː.baɪ.tɐ)]. I would argue, however, that this rendition is only valid under a strong pragmatic pressure to clarify the morphological structure (e.g. in the case of a misunderstanding or correction), and uncommon in running speech.

4 Schlüter (2005) notes the same inversion of the indefinite determiner and the adverb *quite* in English and argues that this inversion has a rhythmic motivation.

b. *ein dermaßen großer Bub*
b.' **dermaßen ein großer Bub*
b." **ein dermaßen ein großer Bub*
‘(a) such (a) big boy’

The ungrammaticality of inversion or doubling in (5) casts doubt on purely syntactic accounts of this phenomenon and instead provides further evidence for a rhythmic trigger for these word order options.

As noted above, apart from stress clashes, sequences of unstressed syllables are considered disrhythmic and are thus avoided.[5] For instance, when the productive suffix *-er* is attached to trochaic place names ending in *-en* [ən] to derive a demonym to the place name, elision of a reduced syllable is common in certain dialects.

The elision of a reduced syllable in these dialects is probably motivated by linguistic rhythm, specifically to avoid sequences of two reduced syllables (*LAPSE). This process seems to have an areal distribution such that it does not affect all place names in the same way, as may be observed when comparing (6) and (7) with (8).[6]

(6) semi-transparent, with resyllabification of stem-final consonant(s) (mainly East Central German and Bavarian)
a. *Dresden* [dʁeːs.dən] – *Dresdner* [dʁeːs.dnɐ]
b. *Bautzen* [baʊ.tsən] – *Bautzner* [baʊ.tsnɐ]
c. *München* [mʏn.çən] – *Münchner* [mʏn.çnɐ]
d. *Weiden* [vaɪ.dən] – *Weidner* [vaɪ.dnɐ]

(7) opaque, elision of stem-final consonant (Northern Low Saxon)
a. *Emden* [ʔɛm.dən] – *Emder* [ʔɛm.dɐ]
b. *Bremen* [bʁeː.mən] – *Bremer* [bʁeː.mɐ]
c. *Norden* [nɔɐ.dən] – *Norder* [nɔɐ.dɐ]
d. *Apen* [ʔaː.pən] – *Aper* [ʔaː.pɐ]

5 The Strict Layer Hypothesis assumed in prosodic phonology (Selkirk 1984) provides a supplementary explanation for the avoidance of lapses: under the assumption that feet in German are trochaic and maximally disyllabic, further unstressed syllables cannot be parsed into feet and thus constitute a violation of the principle EXHAUSTIVITY.

6 The examples represent the written norm within the dialectal areas. For certain place names, reduced and full form coexist, e.g. *Uelzen – Uelzener ~ Uelzer*; *Dülmen – Dülmener ~ Dülmer*.

(8) fully transparent, no effect of *LAPSE
 a. *Hagen* [haː.gən] – *Hagener* [haː.gə.nɐ]
 b. *Siegen* [ziː.gən] – *Siegener* [ziː.gə.nɐ]
 c. *Gießen* [giː.sən] – *Gießener* [giː.sə.nɐ]
 d. *Aachen* [ʔaː.xən] – *Aachener* [ʔaː.xə.nɐ]

The effect of *LAPSE is especially remarkable when considering dactylic place names like *Tübingen*, *Kaufungen*, *Bevensen*. Mere affixation of the demonymic suffix would lead to three consecutive unstressed syllables, a configuration that is ungrammatical across dialects (**Tübingener*, **Kaufungener*, **Bevensener*). Instead, *LAPSE dictates haplology in these cases, resulting in elision of stem-final [ən] – as in (7) – to yield *Tübinger*, *Kaufunger*, *Bevenser*.

Vogel et al. (2015) have shown clear effects of *LAPSE on the linearization of constituents in German sentences. Specifically, Vogel et al. (2015) investigated i) the ordering of inherently weak pronominal adverbs in the Middlefield and ii) auxiliary verbs in sentence final verb clusters. Speakers were to repeat sentences with these constructions that were presented in either a rhythmically alternating or a disrhythmic condition, i.e. one in which the placement of the weak pronoun or auxiliary verb leads to three consecutive, unstressed syllables. The results reveal a clear effect of rhythm such that recall errors were significantly more likely in the disrhythmic conditions. In other words, the syntactic representation of the sentences to be recalled was more stable when the corresponding prosodic representation was rhythmically optimal.

The preceding examples attest the importance of rhythmic alternation, more specifically, of the constraints *CLASH and *LAPSE for the phonological representation (and processing) not only of words but also at the phrasal level. The low level constraints fostering an alternating rhythm of strong and weak syllables are complemented by a tendency for iterating rhythm such that patterns that emerge from the distribution of prominences are preferably repeated.

1.3 Prosodic parallelism within and beyond the word

Recently, Wiese and Speyer (2015) suggested that prosodic parallelism is relevant for the occurrence of final schwa in cases like (1) (see Kentner 2015, for discussion). In a nutshell, the idea is as follows: when given the choice, speakers strive for prosodic parallelism; for two words that are prosodic phrase mates, the foot structures are preferably parallel, i.e. the feet display the same number of syllables and stress pattern. Thus, their argument goes, the appearance or lack of optional schwa is dependent on the foot structure of neighboring words.

Examining a large corpus of written German, Wiese and Speyer (2015) investigated, inter alia, several cases of nouns with apparently freely alternating monosyllabic and disyllabic variants like *Tür* – *Türe* ('door') or *Tags* – *Tages* ('day-GEN') in the context of (preceding) monosyllabic or disyllabic determiners.

(9) a. $((\text{die})_{\Sigma}\ (\text{Tür})_{\Sigma})_{\varphi}$
    b. $((\text{ei.ne})_{\Sigma}\ (\text{Tü.re})_{\Sigma})_{\varphi}$

Using chi-square tests on bigram frequencies, they disprove statistical independence of the prosodic shapes of co-occurring determiner and noun. In a follow-up study, Wiese (2016) reports corroborating evidence in corpora of spontaneous spoken language. These results suggest that, when possible, the prosodic structure of the noun preferably mirrors the structure of the determiner, cf. (9). Note that this explanation assumes that function words like determiners project a foot (see Kentner 2015, for discussion).

The effect of prosodic parallelism is not confined to German schwa-zero alternations alone. In fact, there are phenomena that would defy proper analysis without recourse to a constraint on prosodic parallelism; these are cases in which the PARALLELISM constraint appears to have a stronger influence compared to the German schwa-zero alternations, in which parallelism is merely a tendency. Consider Standard Chinese, in which the productivity of N+N compounds and V+Obj combinations is strictly constrained by the number of syllables. As Duanmu (2012) shows, parallel prosodic structures with either two monosyllables (1+1) or two disyllables (2+2) are generally licit for both constructions. However, for N+N compounds, non-parallel structures of the 1+2 type are mostly unacceptable. Similarly, for V+Obj phrases, the imbalanced pattern 2+1 is considered unacceptable (cf. Luo and Zhou 2010; Luo, Duan, and Zhou 2015, for pertinent neuro- and psycholinguistic evidence).

Another case demonstrating the influence of PARALLELISM, again in German morphophonology, is rhyme and ablaut reduplication (Kentner 2017). This type of reduplication has a strict non-identity requirement concerning base and reduplicant, both of which correspond to a prosodic foot (*schickimicki*, **schickischicki* < *schick* 'posh'; *hickhack*, **hackhack* < *hacken* 'to chop, to bicker'). Crucially, nonidentity is confined to the segmental tier. That is, a difference between base and reduplicant concerning the prosodic shape is prohibited (**schischicki*,

**schickischick*; **hickhacke*, **hickehack*), and it is this prohibition that strongly suggests the workings of prosodic parallelism.

Wiese and Speyer's proposal on prosodic parallelism is in line with the observation that equal-sized prosodic constituents are preferred on various levels of prosodic representation and processing. This finding has been codified in several ways: for instance, Ghini (1993) suggests that, in Italian, prosodic structure is built in a fashion that guarantees balanced phonological phrases even if the resulting phrasing is non-isomorphic to syntactic structure. Similarly, Myrberg (2013), examining Stockholm Swedish, suggests that prosodic subconstituents conjoined within an intonational phrase preferably have the same prosodic status. Selkirk (2000) invokes the constraints BINMIN and BINMAX which jointly favor minimally and maximally two prosodic words per phrase. Féry and Kentner (2010) and Kentner and Féry (2013) propose a Similarity condition on prosodic structure such that neighboring constituents at the same level of syntactic embedding be adjusted to exhibit a similar prosodic rendering, irrespective of the constituents' inherent complexity.

Given the abundance of evidence for parallelism, it is not far-fetched to consider it a well-formedness condition on prosodic structure, just like *CLASH and *LAPSE. The exact formulation of this constraint, however, is open to debate (as is the formulation of *CLASH and *LAPSE, cf. discussions in Anttila, Adams, and Speriosu 2010; Shih et al. 2015). Suffice it to say that the PARALLEL constraint requires adjacent prosodic constituents grouped within a higher constituent to exhibit the same prosodic structure.

Having introduced the three rhythmic constraints and their workings in various environments, the following section assesses their relative contribution to word prosodic structure and phrasal rhythm.

2 Three studies on word prosodic structure and phrasal rhythm

The three studies to be presented below were designed to explore the influence of the rhythmic environment on morphophonological (and morphosyntactic) variation in German. Although prosody (or particularly prosodic rhythm) is not explicitly encoded in the written modality (but see Evertz and Primus 2013), all three experiments use written material for this purpose. This is justified by numerous findings which converge to suggest that the use of the written modality (reading and writing) involves recourse to prosodic representations (see, e.g.

Chafe 1988; Breen 2014; and the collection of studies in Kentner and Steinhauer 2017).

The experiments use different linguistic environments and employ different methodologies but all share as crucial factor the distribution of lexically strong or stressed syllables around the morphophonologically or morphosyntactically variable word(s). Experiment One is an oral reading experiment that focuses on the prosodic rendering of the variable prosodic adverb *gern(e)* in different rhythmic contexts. A large-scale online survey (>150 participants), Experiment Two explores rhythmic influences on the choice between possible realisations of possessive or partitive relations. In Experiment Three, we return to prosodically variable adverbs. Employing a corpus analysis, we investigate the usage frequency of prosodically variable adverb-verb sequences to specifically pit effects of rhythmic alternation (avoidance of stress clash and lapse) against those of prosodic parallelism.

2.1 Rhythmic context effects on optional schwa in read speech

The first experiment is concerned with the effects of the rhythmic-prosodic context on the realisation of the prosodically variable adverb *gern(e)* ('happily') in spontaneous (unprepared) oral reading. This adverb has two graphemic representations that correspond to i) a monosyllabic <*gern*> or ii) a trochaic variant <*gerne*>. For the experiment, both graphemic variants were embedded in sentences with systematically varied rhythmic-prosodic structures to ascertain the effect of the rhythmic context on the realisation of schwa on the adverb in scripted speech production.

Previous work suggests that optional schwa syllables are used by speakers to optimise the rhythmicity of phrases and sentences; specifically, it has been argued that a schwa syllable may act as a buffer syllable that prohibits stress clash (Kuijpers and van Donselaar 1998; Rohdenburg 2014; Schlüter 2005). In the case of the variable adverb <*gern(e)*>, the optional schwa syllable may thus prevent a potential clash with a word to the right of it.

The first manipulation of this experiment therefore targets the syllable to the right of the variable word: the noun following the variable adverb in (10) begins in either a stressed (*Himbeeren*) or an unstressed syllable (*Kartoffeln*). In addition, the rhythmic context to the left of the word was manipulated; this manipulation is motivated by the hypothesised propensity for iterating or sequential rhythm that is at the core of prosodic parallelism. The lexical material of the sentences was constructed to yield a trochaic beat with every other syllable bearing lexical stress. The syllabic structure of the noun directly preceding the target adverb <*gern(e)*> was systematically varied, with either a monosyllable (*Hof*) or a disyllabic trochee (*Garten*). Thus, the first (or only) syllable of the variable adverb falls on either an on-beat or off-beat position of the established trochaic pattern. According to the parallelism hypothesis, the trochaic form of the variable adverb should be preferred when preceded by a trochee while the monosyllabic form should be preferred when preceded by a monosyllabic foot.

(10) a. *Bodo will in Steffis Garten gerne Himbeeren ernten.*
b. *Bodo will in Steffis Garten gern Himbeeren ernten.*
c. *Bodo will in Steffis Hof gerne Himbeeren ernten.*
d. *Bodo will in Steffis Hof gern Himbeeren ernten.*
e. *Bodo will in Steffis Garten gerne Kartoffeln ernten.*
f. *Bodo will in Steffis Garten gern Kartoffeln ernten.*
g. *Bodo will in Steffis Hof gerne Kartoffeln ernten.*
h. *Bodo will in Steffis Hof gern Kartoffeln ernten.*
'Bodo would like to harvest {raspberries, potatoes} in Steffi's {yard, garden}'

2.1.1 Materials, participants, procedure

Twenty-four item sets like (10) were devised. The items were distributed over eight lists such that items and conditions were counterbalanced across the lists with each list containing exactly one condition from each item set. Additionally, each list contained 64 filler items from four unrelated experiments and three practice items not connected to any of the experimental items, yielding a total of 91 items. With the exception of the three initial practice items, the item order was determined by pseudo-randomization (van Casteren and Davis 2006) (for each participant individually) such that items from the same experiment had a minimal distance of two intervening items from other experiments and items from the same experimental condition were separated by at least three fillers.

Twenty-four members (19 female) of the Goethe-University community (Frankfurt, Germany) took part in the experiment. All participants are native speakers of German with normal or corrected-to-normal vision per self report. Initially, participants were not informed about the purpose of the experiment but debriefed after the experiment ended.

The experiment took place in a silent office at Goethe University in single sessions for each participant. Participants were seated in front of a 21.5-inch computer screen and equipped with a microphone head set (Shure) attached to an R-44 digital recorder.

All 91 items of each list were presented in a slide show. Each item was presented on two consecutive screen displays. The first display presented two (irrelevant) context sentences in the upper half and the first two words of the target sentence (in the case of this experiment: subject and modal verb) in the middle of the screen (all text left-aligned). Upon pressing the enter button on the keyboard, the target sentence appeared in full (leaving the rest of the first display intact). Participants were asked to read the first display (i.e. the context) silently before moving on to the second display screen. To ensure spontaneous, unprepared oral reading and minimal look-ahead, participants were instructed to read out the target sentence immediately as it appeared on screen and to do so as fluently as possible. The participants were discouraged from making corrections during or after reading and to move on to the next item after reading by another button press. The productions of the participants were recorded on a digital memory card.

2.1.2 Results

All in all, (24 items x 24 participants =) 576 experimental sentences were recorded. Two student assistants independently evaluated each target sentence. Their task was to determine by ear i) whether the production was a fluent and flawless response to the target sentence and ii) whether the target adverb was realised as monosyllablic *gern* or disyllabic *gerne*.

Seven sentences (1.2%) were scored as non-fluent or otherwise flawed by at least one referee and discarded from further analysis. The judgments concerning the number of syllables were perfectly consistent. Aggregating the 569 valid responses, the adverb was judged to contain a reduced syllable in 260 cases (45.7%) and monosyllabic in 309 cases (54.3%) suggesting a slight preference for the monosyllabic form. All in all, the oral realisation of the adverb corresponded to the graphemic representation in 82% of the cases.

Mixed logistic models (Bates et al. 2013) were applied in the statistical computing environment R (R Core Team 2015) to assess the effects of the graphemic representation (‘writtenE’), the rhythmic environment to the left (‘RhythmLeft’), and the rhythmic environment to the right (‘RhythmRight’) on the realisation of the schwa syllable (dependent variable: ‘realiseE’) in reading. The fixed effects (or predictor variables) were coded as orthogonal sum contrasts to ensure minimal correlation. Apart from the fixed effects, the models included participant ('speaker') and item as random effects that were adjusted for by participant or by item differences in the effects of the predictor variables. Likelihood ratio tests (carried out by the anova function) were used to compare models with different predictor variables and random effect structures in order to determine the model with the best fit for the data. The likelihood ratio test generally prefers simpler or more parsimonious models over more complex ones as long as the inclusion of model parameters does not significantly increase model fit. Consequentially, if the inclusion of a model parameter did not significantly improve model fit, it was culled from the model. Complex models with all three predictor variables, the respective interactions, and complex random effects structures[7] were tested first and non-significant predictors (as determined by the likelihood ratio tests) were culled in a stepwise fashion. Over and above a highly significant effect of the graphemic representation (readers preferably realise the adverb in line with its graphemic representation), the preceding context significantly affects the realisation of the reduced syllable. As visible in Table 1, trochaic *gerne* appears to be more likely when the preceding word is trochaic; conversely, the monosyllabic variant is preferred after monosyllabic nouns. The coefficients of the best fitting logistic mixed model are tabulated in Table 2.

Tab. 1: Percentages of trochaic realisations of the variable adverb *gern(e)* broken down by graphemic form of stimulus and prosodic form of preceding noun

Percentage of trochaic realisations of adverb		**Prosodic form of noun preceding the adverb**	
		trochaic	**monosyllabic**
Graphemic form of the adverb	<gern>	17	10
	<gerne>	80	76

7 Several of the more complex models did not converge. Non-converging models were not considered further in the model comparison process.

Tab. 2: Coefficients of the best fitting mixed logistic model with the formula *glmer(realiseE~writtenE+RhythmL+(writtenE | speaker)*, family = binomial). N = 569

	Estimate	Std. Error	z-value	p-value
(Intercept)	−0.06767	0.43364	−0.156	0.8760
writtenE	2.50456	0.33543	7.467	<0.001
RhythmLeft	−0.27643	0.13822	−2.00	0.0455

2.1.3 Discussion

The experiment shows that readers are generally guided by the written form of the prosodically variable word when producing it in spontaneous read speech. Apart from the effect of the graphemic representation, the rhythmic context has a small but significant effect on the realisation of schwa on the critical adverb. This, however, only holds for the manipulation concerning the rhythmic structure to the left of the critical word (RhythmLeft). The other rhythmic effect that was tested in this experiment, the rhythmic context to the right of the critical word, failed to affect the realisation of *gern(e)*. One conceivable explanation is related to the task of spontaneous oral reading: readers may simply not have had the time to sufficiently process the upcoming word to prosodically adjust the target word to it.

The significant effect of RhythmLeft suggests that readers prefer monosyllabic *gern* after a monosyllabic noun while trochaic *gerne* preferentially follows a trochaic noun. This finding, at first sight, corroborates the prediction according to the parallelism hypothesis. However, taking into account the wider prosodic context (with the trochaic beat that was established right from the beginning of the sentence), parallelism as formulated by Wiese and Speyer (2015) may be insufficient to explain the results. Under Wiese and Speyer's account, and under the Strict Layer Hypothesis (SLH) of prosodic phonology (Selkirk 1984), feet cannot straddle word boundaries. This limitation, however, is crucial when evaluating the parallelism effect. Compare, in this respect, the conflicting footings of an example item in (11):

(11) Conceivable foot structures
  a. trochaic footing ('Abercrombian' feet)
    *(Rosie) (will auf) (jeden) (**Fall gern**) (Ärztin) (werden)*

b. footing according to the Strict Layer Hypothesis
*(Rosie) (will) (auf) (jeden) (**Fall) (gern**) (Ärztin) (werden)*
Rosie wants in any case happily physician become
'In any case, Rosie would like to become a physician.'

(11a.) represents a perfectly iterating prosodic structure – a sequence of six trochees – but blatantly violates the Strict Layer Hypothesis, with the adverb *gern* demoted to the weak position of a trochee (in this position presumably being unaccentable); (11b.), in contrast, abides by the letter of the SLH but the structure fails to represent the trochaic beat that is felt when the sentence is uttered. This is because, according to Wiese and Speyer (2015), even function words project feet (see Kentner 2015, for criticism). If one were to follow Wiese and Speyer (2015), the intended trochaic beat of the experimental items does not correspond to parallel prosodic structures in the first place.

The analysis of the read sentences so far only considered the presence or absence of schwa on the critical adverb but did not involve any assessment of its prosodic prominence. A cursory look at the realisations of the adverb, however, suggests that the monosyllabic adverb often remains entirely unaccented (which would be in line with the representation in (11a.)) and often features a centralised vowel: [gɐn]. There is independent evidence to the effect that leaving the adverb unaccented (a necessity for the representation in (11a.)) is very common: Kutscher (2014) found that adverbs in German are often prosodically reduced, and thus serve as a trough between prominence peaks, preventing stress clash.

While I acknowledge that this experiment cannot settle the largely theoretical debate among the schools favoring Abercrombian feet (11a.) over those abiding by the SLH (11b.) or vice versa, I point out that the representation (11a.) not only respects *Clash and *Lapse; (11a.) also exhibits a sequential rhythm and may thus be in line with a weaker version of Parallelism that tolerates violations of the SLH. (11b.), in contrast, only locally fulfills the Parallelism constraint (in the bolded part of the sentence) but fails to respect other constraints on rhythmic structure (*Clash, *Lapse) in spite of the fact that a natural rendition of the sentence exhibits a perfect alternation of prominences.

2.2 Rhythm and morphosyntactic choice: Morphological genitive vs. prepositional construction

In German, the possessive or partitive relationship may be expressed by (at least) two syntactically distinct constructions:[8] by morphological case (genitive) or by a prepositional phrase headed by *von* ('of'). The choice between these two is partly governed by register or style with the prepositional construction deemed more colloquial and the genitive more formal. Given that the prepositional construction affords more (function) words than the morphological genitive, the two variants also exhibit a difference concerning their rhythmic patterns.

(12) a. *Der Wirt der Herberge*
the owner the.GEN inn.GEN
b. *Der Wirt von der Herberge*
the owner of the.DAT inn.DAT
'the owner of the inn'

As apparent from a comparison of the two syntactic options in (12), the prepositional phrase (12b.) involves, in addition to the determiner, a (usually) unaccented syllable (the preposition *von*) which increases the distance between head noun and attribute. When the latter two are lexical words and new to the discourse context, these referents usually bear an accent. The exact location of the two accents and their distance from each other depends not only on the construction (genitive or prepositional phrase) but also on the prosodic structures of the lexical words involved. The accents are the further apart the more unstressed syllables follow the head noun's stressed syllable, or the more unstressed syllables the attributive noun has preceding its stressed one.

Making use of a systematic manipulation of the prosodic distance between head noun accent and accent on the attribute, the following study aims at testing the hypothesis that the choice between the two syntactic options is attributable to the (implicit) rhythmic structure they engender. A very similar hypothesis has recently been confirmed for the usage of the English *s*-genitive and 'of'-genitive by Shih et al. (2015) who conducted a large-scale analysis using a corpus of spoken American English; in their dataset, however, the effects of rhythm on construction choice, although detectable, were largely dampened by the factor animacy. Here, a more controlled experimental avenue was chosen, i.e. a question-

8 In the following, further options will be disregarded, e.g. compounding *Herbergswirt* (lit. 'inn owner') or the preposed genitive, as in *Marias Hund* ('Maria's dog'). The latter construction is confined to animate genitives and mainly used with proper names.

naire study in which the prosodic structures of both the head noun and the attribute were systematically varied while leaving the factor animacy constant. The study will be detailed in the following.

2.2.1 Materials and method

An online questionnaire (Sosci Survey by Leiner 2014) was set up in which participants had to tick their preferred option for the expression of a possessive or partitive relationship in various rhythmic conditions. To this end, 24 items like (13a.–d.) were devised with head noun and attribute separated by a blank. The four conditions of the 24 items were counterbalanced across four blocks in a latin square design such that no head noun and attribute was presented more than once per block. The trials were presented in randomised order, interspersed with 40 filler items from two unrelated experiments. Each item was presented on a single slide together with four options to fill the blank. By ticking the appropriate box, participants had to choose either *der* (i.e. the monosyllabic definite determiner for the genitive attribute) or *einer* (i.e. the disyllabic indefinite determiner for the genitive attribute) or *von der* (i.e. the preposition and following determiner for the prepositional construction). In addition, a fourth option (*aus* 'from') was given as an oddball option that invariably leads to an ungrammatical construction. This was included to be able to spot participants who randomly marked one option without proper consideration of the item.

175 students of the Goethe-University community participated in the online questionnaire. Each participant was randomly assigned to one block.

(13) Insert *der* or *einer* or *von der* or *aus*
   a. *Der Knopf ... Arbeitshose*[9]
   b. *Die Knöpfe ... Arbeitshose*
   c. *Der Knopf ... Gesäßtasche*
   d. *Die Knöpfe ... Gesäßtasche*
   'the button(s) {a., c.: Sg; b., d.: Pl} of the {a., b.: work pants; c., d.: back pocket}'

9 All attribute nouns have feminine gender in order to avoid fusion of preposition and determiner, a common process with masculine or neuter attributes (*von dem* > *vom*).

2.2.2 Predictions

The study was originally designed to test the hypothesis that construction choice is co-determined by the propensity for rhythmic alternation. Correspondingly, more short genitives (the monosyllabic determiner *der*) are expected in conditions with greater distance between the accented syllables of head noun and attribute, i.e. when the head noun has non-final stress and the attribute has non-initial stress. Conversely, more prepositional constructions *von der* or disyllabic genitives *einer* are expected in conditions with a short distance between the accented syllables of head noun and attribute. Opposing predictions come about when considering effects of prosodic parallelism. According to the PARALLELISM constraint, structures are preferred that yield an iterating rhythm. Correspondingly, in our case, a trochaic head noun (such as *Knöpfe*) should give rise to a preference for the disyllabic determiner *einer* or the prepositional construction with *von der* (the monosyllabic preposition and monosyllabic determiner are assumed to be grouped into a trochee). A monosyllabic head noun, in turn, should promote the monosyllabic determiner der.[10]

Tab. 3: Percentages for chosen possessive/partitive construction broken down by stress on head noun and attribute

		Ultima of head noun		Initial of attribute noun	
		unstressed	**stressed**	**unstressed**	**stressed**
Prosodic form of Gen or PP	monosyllabic	59	55	58	56
	trochaic	41	45	42	44
	Total	100	100	100	100

10 The prosodic structure of the attribute was varied in such a way as to make predictions according to prosodic parallelism impossible to test with respect to the attribute noun. The first syllable of the attribute was either a stressed syllable or an unstressable reduced syllable. In the latter case it is unclear what kind of material would be preferred, according to parallelism, between head noun and attribute. What is more, the foot structures of the attributes with initial stress was variable, i.e. some items had initial monosyllabic feet (e.g. *(Haupt)(schu.le)*), some presented with trochaic initials (e.g. *(Ei.sen)(bahn)*).

2.2.3 Results

Several participants only partly completed the questionnaire, resulting in many missing answers. All in all, 3662 responses or 87% of the expected 4200 (= 175 participants x 24 items) were collected. In 39 of the cases, the oddball option *aus* was chosen, resulting in ungrammatical constructions. The majority (90%) of the remaining 3623 valid responses resulted in a genitive construction (2055 times or 57% of the cases monosyllabic *der*; 1195 times or 33% of the cases disyllabic *einer*). In only 373 or 10% of the cases, the prepositional construction was chosen. One reason for this discrepancy lies in the fact that there were two options to choose from genitives but only one valid prepositional option (not counting the ungrammatical oddball). Moreover, since the task was presented in writing, there is certainly a tendency to choose the formal genitive over the more colloquial prepositional construction.

In order to specifically test the predictions according to the principle of rhythmic alternation (*Clash, *Lapse) and Parallelism, the responses were grouped by prosodic structure, i.e. the disyllabic trochaic genitive determiner *einer* was collapsed with the likewise trochaic prepositional *von der* and juxtaposed to the monosyllabic genitive determiner *der*.

Table 3 shows the percentages of monosyllabic (*der*) vs. disyllabic responses (*einer* or *von der*) broken down by the prosodic status (stressed or unstressed) of the ultima of the head noun and the initial syllable of the attribute noun. Clearly, participants gave more disyllabic responses when the head noun presented with stress on the ultima and when the attribute had initial stress.

Logistic mixed models (Bates et al. 2013) were applied to assess the effects of the prosodic status of the head noun (stressed or unstressed ultima) as well as of the attribute noun (stressed or unstressed initial syllable) on the choice of monosyllabic or disyllabic responses. The intercepts for participants and items were included as random effects. Again, as in the previous study, predictor variables (which were coded as orthogonal sum contrasts) were culled from the model when their inclusion did not improve model fit.

The results of the best fitting logistic mixed model are tabulated in Table 4. Contrary to predictions, including the effect of stress position on the attribute does not improve model fit. However, the model confirms that the prosodic structure of the head noun significantly affects the choice of the construction. With an unstressed ultima on the head noun, the monosyllabic determiner is clearly preferred over the disyllabic genitive or prepositional construction, most likely because the latter would yield a disrhythmic structure with three or four unaccented syllables in a row. The results thus support the hypothesis that participants strive for rhythmic alternation when making syntactic decisions.[11]

There is, however, no indication that participants build prosodically parallel structures. According to prosodic parallelism, as conceived by Wiese and Speyer (2015), participants would have had to prefer a disyllabic trochaic genitive or preposition plus determiner after a trochaic head noun, or, conversely, a monosyllabic genitive after a head noun featuring a stressed ultima. This is clearly not the case.

To conclude, the propensity for rhythmic alternation has a significant impact on construction choice while any effect of prosodic parallelism remains mute.

Tab. 4: Coefficients of the best fitting generalised linear model evaluating the choice of the possessive/partitive construction

	Estimate	**Std. Error**	**z-value**	**p-value**
(Intercept)	−0.33166	0.14641	−2.265	0.0235
UltimaHeadNoun	0.09413	0.03711	2.536	0.0112

2.3 *CLASH, *LAPSE, PARALLELISM – a corpus study

The third study examines a large-scale corpus (DeReKo, cf. Institut für Deutsche Sprache [IDS]) to directly compare the effects of *CLASH, *LAPSE, and PARALLELISM on the morphophonological variation concerning the German adverbs *gern* ~ *gerne* ('happily'), *lang* ~ *lange* ('for long'), *selbst* ~ *selber* ('{my-, your-, her-, him-, our-, them}-{self-, -ves}'), and *meist* ~ *meistens* ('most of the time'). These adverbs display a (free) alternation concerning the schwa and, consequentially, syllabic structure, i.e. they feature either a monosyllable or a trochee. In the latter case, the final syllable is always a reduced syllable (schwa or [ɐ] in the case of *selber*). Importantly, in contrast to further alternating adverbs, the allomorphs of these adverbs have graphemic cognates both of which are equally acceptable in written Standard German. To the best of my knowledge, there are no more alternating adverbs both variants of which are likewise acceptable in writing.

11 In recent years, a number of online experiments studying eye movements in reading showed comparable results which suggest that the rhythmic/prosodic environment affects syntactic parsing decisions in written sentence comprehension (Breen and Clifton 2013; Kentner 2012; Kentner and Vasishth 2016).

2.3.1 Method and materials

The frequencies of the four variable adverbs were examined in the context of two forms of the verbs *tun* and *machen* ('to do', 'to make') when these follow the variable adverb. This way, four variable adverbs by two verb forms, i.e., eight quadruplets of prosodically different adverb-verb combinations were scrutinised.

Tab. 5: Bigrams scrutinised in corpus experiment and corresponding factors used for the evaluation of the rhythmic effects

Adverb	Verb	*CLASH	*LAPSE	PARALLEL
gern/selbst/meist/lang	*tun*	×	✓	✓
gerne/selber/meistens/lange	*tun*	✓	✓	×
gern/selbst/meist/lang	*getan*	✓	✓	×
gerne/selber/meistens/lange	*getan*	✓	×	×
gern/selbst/meist/lang	*machen*	×	✓	×
gerne/selber/meistens/lange	*machen*	✓	✓	✓
gern/selbst/meist/lang	*gemacht*	✓	✓	×
gerne/selber/meistens/lange	*gemacht*	✓	×	×

The prosodic profile of each bigram was coded according to the three rhythmic constraints. This was done in a binary fashion, as displayed in Table 5, where the bigrams are represented as either respecting or violating each of the three constraints respectively.

For each of the four combinations of verb form and adverb, the bigram frequencies within the DeReKo corpus, written section (Institut für Deutsche Sprache [IDS]) were determined. Chi-square tests were applied to test the statistical independence of adverb and verb form. These tests use contingency tables like (14) to compare the expected frequencies according to the null hypothesis (which assumes adverb and verb forms to be statistically independent from each other) to the actual, observed frequencies.

(14)

		Prosodic form of verb	
		Monosyllabic	**iambic**
Prosodic form of adverb	monosyllabic	*gern tun*	*gern getan*
	trochaic	*gerne tun*	*gerne getan*

2.3.2 Data analysis and results

For seven of the eight quadruplets of bigrams, the Chi-square tests clearly disprove statistical independence of the prosodic structure of adverb and verb (with p-values < 0.01), supporting the hypothesis that the choice of the prosodic form is conditioned by the prosodic shape of the context. Only in the case of the *meist(ens) machen/gemacht* quadruplet, the test did not yield a significant result. In any case, it has to be determined whether and to what extent each of the three rhythmic constraints under discussion contribute to the prosodic effect. Therefore, for each of the 32 bigrams, the standardised Chi square residuals[12] were calculated as a measure for the degree of deviance from assumed statistical independence of the prosodic form of the adverb and the prosodic form of the verb. Testing the predictions of the three rhythmic constraints against the residuals can inform us about the extent to which each constraint contributes to the frequency distribution of the adverb-verb combinations. In general, a negative residual indicates that a bigram occurs less frequently than the null hypothesis would lead one to expect; conversely, a positive value indicates that the bigram is used more frequently than expected. That is, if the constraints were to affect the prosodic form of the adverb-verb bigrams, structures that violate a given constraint should obtain negative residuals, while bigrams that respect the constraint should engender more positive residuals.

12 Standardised residuals are calculated as (Observed Frequency – Expected Frequency) / sqrt(Expected Frequency)

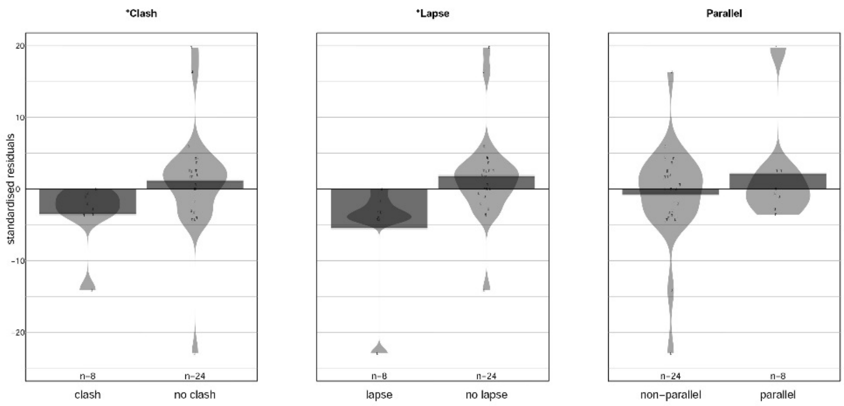


Fig. 1: Standardised Chi square residuals (y-axis) broken down by the two levels of each factor, representing the constraints *CLASH (left panel), *LAPSE (middle panel) and PARALLELISM (right panel), respectively (x-axis). The dots correspond to the individual residuals for each bigram (n=32). The bars represent the mean residual, and the shaded area around the dots illustrates the density of the distribution (the wider the shaded area, the denser the clustering of the residuals in that area).

In order to get a first impression about the contribution of the three constraints, the 32 standardised residuals are plotted for each level of the three predictor variables using the YaRrr package (Phillips 2017) in the statistical computing environment R (R Core Team 2015). As the plot in Figure 1 shows, bigrams that violate a constraint are, on average, less frequent than expected according to the null hypothesis and hence show more negative residuals (cf. left bars of the three panels) while bigrams respecting the constraints are more frequent than expected. The distribution of residuals thus corroborates the hypothesis that adverb-verb bigrams that respect the rhythmic constraints are favored over those bigrams that violate the relevant constraints. However, the differences between the residuals for the bigrams that violate versus bigrams that obey a given constraint are clearly more pronounced in the case of *CLASH and *LAPSE than in the case of PARALLELISM. This is especially apparent in the residuals for those bigrams that violate *CLASH and *LAPSE: Almost all residuals for bigrams that involve a clash or a lapse are negative, while the residuals for the non-parallel bigrams (left bar in the right panel of Fig. 1) are more evenly distributed with the mean residual close to zero. The distribution of residuals corresponding to the bigrams respecting *CLASH or *LAPSE (right bars in the left and middle panel), while positive on average, spans both the positive and the negative range (most likely due to the fact that bigrams that obey *CLASH may violate *LAPSE, and vice versa).

Linear models (Bates et al. 2013) were employed to analyse the data. The standardised residuals that were calculated for each of the 32 bigrams (see above) were used as dependent variable. The three constraints (*CLASH, *LAPSE, PARALLELISM) served as binary predictor variables, with each bigram violating or respecting the constraints (cf. Table 3); these predictors were coded as orthogonal contrasts. Including the specific adverb as grouping variable (random effect) did not improve model fit. In Table 6, the output of the model including all three predictor variables is tabulated, with *CLASH and *LAPSE clearly showing significant effects while the effect of PARALLELISM remains non-significant.

A second, simpler model was fit with PARALLELISM discarded as predictor (cf. Table 7). Applying the anova function to compare the simpler model with the full model suggests that discarding PARALLELISM does not deteriorate model fit (Df = 1, p = 0.32).

To summarize, the negative Chi square residuals for bigrams involving a stress clash (e.g. /gern machen/) or a stress lapse (e.g. /gerne getan/) reflect the avoidance of these rhythmically sub-optimal structures when compared to bigrams that obey the respective constraints. No such pattern of avoidance could be observed for bigrams that violate the PARALLELISM constraint (i.e. non-parallel bigrams like /gerne tun/ or /gern machen/). This corpus study thus corroborates the hypothesis that the inclusion or omission of the optional schwa-syllable on the adverb is conditioned by the stress status of the initial syllable of the verb. The overall prosodic shape of the verb, however, i.e. whether it is monosyllabic, trochaic, or iambic, does not appear to affect the inclusion/omission of the schwa syllable on the adverb beyond the effects of *CLASH and *LAPSE.

Tab. 6: Model including all three main effects

	Estimate	Std. Error	t-value	p-value
(Intercept)	−0.02698	1.02653	−0.026	0.979212
*Clash	2.13	0.64902	3.282	0.00277
*Lapse	2.29243	0.64902	3.532	0.00145
Parallelism	0.65506	0.64902	1.009	0.32148

Tab. 7: Model with main effect of PARALLELISM culled

	Estimate	Std. Error	t-value	p-value
(Intercept)	−0.02698	1.02653	−0.026	0.979212
*Clash	1.96624	0.62862	3.128	0.003987
*Lapse	2.45619	0.62862	3.907	0.000514

2.3.3 Discussion

This corpus study yields important insights regarding the morphophonological variation on the adverbs under study. First of all, provided that the written corpus does in fact reflect prosodic preferences, it is clear from the results that supralexical prosodic structure co-determines the presence or absence of a reduced syllable on the variable adverbs. This is in accordance with similar findings by Ingason (2015), Kaufmann (2014), Schlüter (2005), and Vogel et al. (2015) who report rhythmic influences on morphological or morphosyntactic variation. Secondly, this study fails to replicate the findings by Wiese and Speyer (2015) who hold prosodic parallelism accountable for the presence or absence of a reduced syllable. In this study, PARALLELISM does not appear to contribute to the morphophonological variation of the adverbs. The model comparison suggests that the rhythmic influences are reducible to *CLASH and *LAPSE alone. One conceivable reason for the discrepancy between the present results and the findings by Wiese and Speyer (2015) lies in the difference between the structures scrutinised: while this study looked at prosodically variable adverb-verb sequences (e.g. *gern(e) tun*), Wiese and Speyer (2015) studied the variable adverb in other contexts (e.g. the verb-adverb sequence *wär(e) gern(e)*). It remains to be seen why prosodic parallelism explains the variation in one case but not in the other. In this context, it would also be interesting to check to what extend the rhythmic constraints *CLASH and *LAPSE contribute to the variance in Wiese and Speyer's dataset.

Furthermore, this study reveals an interesting finding regarding the relative contributions of *CLASH and *LAPSE, with the latter apparently having a similar, if not stronger, impact on morphophonological choice when compared to *CLASH. Given the greater attention to stress clash and its avoidance in the literature and the comparatively limited consideration of the *LAPSE constraint, this may seem astonishing (entering the terms 'clash' and 'lapse' in the context of the phrase 'linguistic rhythm' produces 493 hits for 'clash' but only 271 for 'lapse' on Google Scholar). What is more, as noted by Julia Schlüter,

> [...] many authors [...] concur in the view that stress clashes are perceived as far more objectionable than stress lapses; while the latter are tolerated to a certain extent, the former almost categorically necessitate compensatory measures.
>
> (Schlüter 2005, 20)

Possibly, the somewhat weaker effect of *CLASH on presence or absence of schwa is due to the fact that a stress clash may be alleviated in other ways, e.g. by stress retraction or stress promotion, processes that the writer may subconsciously execute (remember that we are dealing with data from a written corpus). Conversely, it is hardly possible to change a structure violating *LAPSE by altering the assignment of prominences to syllables because the unstressable reduced syllables simply cannot become stressed. A writer abiding by the principle of rhythmic alternation is thus more likely to put morphophonological variation to its rhythmic use in the event of a potential lapse than in the event of a potential clash (see Shih et al. 2015, for a similar point).

In the following, I note several limitations of this study. For one thing, since I examined the variable structures within a written corpus only, it remains unclear whether the results are generalisable to the oral modality. Even more importantly, since only bigrams were studied, with the wider (prosodic) context disregarded, the validity of the results is open to suspicion. It is quite possible that an analysis that considers the phrasal context would lead to different results. However, the approach taken here is in keeping with Wiese and Speyer (2015) who also only considered bigrams, rendering the studies at least methodologically comparable. Finally, the scope of this study is very narrow, narrower by far compared to Wiese and Speyer (2015) who consider schwa-zero variation not only on adverbs but in many more contexts. The results therefore have to be taken with some caution.

3 General discussion and conclusion

Overall, the three studies presented here clearly support the claim that the rhythmic-prosodic context affects morphophonological variation. The first study revealed an effect of the rhythmic pattern (due to the distribution of lexical stresses) on the realisation of the variable adverb *gern(e)* in oral reading. The second study, a forced choice experiment, showed that the variable morphosyntax of the possessive or partitive relation is susceptible to rhythmic structure. Finally, a corpus study demonstrates the non-independence of the prosodic shapes of variable adverbs in adverb-verb sequences.

As to the relative contribution of the three constraints under discussion (*CLASH, *LAPSE, PARALLELISM) for explaining the variance observed in the three experiments, the findings paint a somewhat mixed picture. The final corpus study quite clearly dismisses the importance of PARALLELISM, while showing that *LAPSE and *CLASH, have a clear impact on the choice of monosyllabic vs. trochaic adverb. Similarly, the experiment on the choice between morphological genitive and prepositional phrase reveals a weak effect of rhythmic alternation but fails to reveal an effect of prosodic parallelism.

The first experiment, however, suggests that prosodic parallelism has a role to play in the realisation of the variable adverb in oral reading. It shows that an iterating rhythm is effectively priming the morphophonological form of the variable adverb that continues the preceding (trochaic) rhythm. However, as highlighted in the discussion of that experiment, the iterating rhythm is only observable through the lens of certain assumptions regarding the foot structure involved, i.e. it is only valid when foot boundaries are allowed to straddle word boundaries (contra the Strict Layer Hypothesis) and when adverbs may be demoted to a prosodically weak position. That is, while there is clear evidence for the joint effects of *CLASH and *LAPSE conditioning the morphophonological structure of words and phrases, effects of prosodic parallelism are relatively minute. This is not to contest the relevance for prosodic parallelism in other contexts. As discussed in the introduction, prosodic parallelism is likely to be a constraining factor in word formation (e.g. reduplication) and it is clearly involved in poetic language. Quite possibly, the role of PARALLELISM is more pronounced in more artistic language use or, more generally in circumstances that are not as strictly constrained by time. Note that for PARALLELISM to become apparent, the linguistic processor needs to consider more material (at least two adjacent feet) than when evaluating local rhythmic well-formedness on a syllable-to-syllable basis.

All in all, the results of the studies presented suggest that phrases and sentences are not built by merely concatenating morphs according to a pre-specified syntactic structure. In addition, word forms may be altered in various ways to suit the supra-lexical rhythmic structure, and the rhythmic structure may reciprocally codetermine morphosyntactic choice.

Acknowledgements: I thank the organisers and the audience of the workshop on empirical approaches to the phonological structure of words (Marburg, October 2014) for the opportunity to present and discuss this work. Caroline Féry, Isabelle Franz, Ingo Plag, Renate Raffelsiefen and two anonymous reviewers provided helpful comments and constructive feedback on the manuscript.

References

Anttila, Arto, Matthew Adams & Michael Speriosu. 2010. The role of prosody in the English dative alternation. Language and Cognitive Processes 25. 946–981.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2013. lme4: Linear mixed-effects models using eigen and s4. R package version 1. https://CRAN.R-project.org/package=lme4 (accessed 12 January 2015).

Behaghel, Otto. 1930. Zur Wortstellung des Deutschen. Language 6(4). 29–33.

Breen, Mara. 2014. Empirical investigations of the role of implicit prosody in sentence processing. Language and Linguistics Compass 8. 37–50.

Breen, Mara & Charles Clifton Jr. 2013. Stress matters revisited: A boundary change experiment. The Quarterly Journal of Experimental Psychology 66. 1896–1909.

Casteren, Maarten van & Matt Davis. 2006. Mix, a program for pseudorandomization. Behavior Research Methods 38. 584–589.

Chafe, Wallace. 1988. Punctuation and the prosody of written language. Written Communication 5. 395–426.

Domahs, Ulrike, Frank Domahs & Christina Kauschke. 2017. The morphology-prosody interface in typically developing and language-impaired populations: An overview and insights from experimental approaches. In Christiane Ulbrich, Alexander Werth & Richard Wiese (eds.), Empirical approaches to the phonological structure of words (Linguistische Arbeiten 567), 91–116. Berlin & Boston: De Gruyter.

Duanmu, San. 2012. Word-length preferences in Chinese: a corpus study. Journal of East Asian Linguistics 21. 89–114.

Eisenberg, Peter. 1991. Syllabische Struktur und Wortakzent: Prinzipien der Prosodik deutscher Wörter. Zeitschrift für Sprachwissenschaft 10. 37–64.

Evertz, Martin & Beatrice Primus. 2013. The graphematic foot in English and German. Writing Systems Research 5. 1–23.

Fanselow, Gisbert & Caroline Féry. 2002. Ineffability in grammar. In Gisbert Fanselow & Caroline Féry (eds.), Resolving conflicts in grammars: Optimality theory in syntax, morphology, and phonology, 265–307. Hamburg: Helmut Buske Verlag.

Fehringer, Carol. 2011. Allomorphy in the German genitive: A paradigmatic account. Zeitschrift für Germanistische Linguistik 39. 90–112.

Féry, Caroline. 1994. Umlaut and Inflection in German. ROA – Rutgers Optimality Archive 34. 1–44.

Féry, Caroline. 1997. Uni und Studis: die besten Wörter des Deutschen. Linguistische Berichte 172. 461–489.

Féry, Caroline & Gerrit Kentner. 2010. The prosody of embedded coordinations in German and Hindi. In Proceedings of Speech Prosody, vol. 5. Chicago.

Ghini, Mirco. 1993. ɸ-formation in Italian: A new proposal. Toronto Working Papers in Linguistics, 41–78.
Gutzmann, Daniel & Katharina Turgay. 2015. Expressive intensifiers and external degree modification. The Journal of Comparative Germanic Linguistics 17. 185–228.
Ingason, Anton Karl. 2015. Rhythmic preferences in morphosyntactic variation and the theory of loser candidates. In Ralf Vogel & Ruben van de Vijver (eds.), Rhythm in Cognition and Grammar: A Germanic Perspective, 235–253. Berlin: De Gruyter.
Institut für Deutsche Sprache (IDS). 2014. Deutsches Referenzkorpus. Archiv der Deutschen Gegenwartssprache [Computer Software]. https://cosmas2.ids-mannheim.de/cosmas2-web/ (accessed 12 January 2015).
Itô, Junko & Armin Mester. 1997. Sympathy theory and German truncations. University of Maryland Working Papers in Linguistics 5. 117–139.
Kager, René 1989. A metrical theory of stress and destressing in English and Dutch. Dordrecht: Foris.
Kallulli, Dalina & Antonia Rothmayr. 2008. The syntax and semantics of indefinite determiner doubling constructions in varieties of German. The Journal of Comparative Germanic Linguistics 11. 95–136.
Kaufmann, Göz. 2014. Die Tür(e) in der deutschen Literatur – Tor zum Verständnis sprachlicher Variation? In Pia Bergmann, Karin Birkner, Peter Gilles, Helmut Spiekermann & Tobias Streck (eds.), Sprache im Gebrauch: räumlich, zeitlich, interaktional. Festschrift für Peter Auer, 123–142. Heidelberg: Winter.
Kentner, Gerrit. 2012. Linguistic rhythm guides parsing decisions in written sentence comprehension. Cognition 123. 1–20.
Kentner, Gerrit. 2015. Problems of prosodic parallelism: A reply to Wiese & Speyer (2015). Linguistics 53. 1233–1241.
Kentner, Gerrit. 2017. On the emergence of reduplication in German morphophonology. Zeitschrift für Sprachwissenschaft 36.
Kentner, Gerrit & Caroline Féry. 2013. A new approach to prosodic grouping. The Linguistic Review 30. 277–311.
Kentner, Gerrit & Karsten Steinhauer. 2017. Prosody in written language – acquisition, processing, emergence of norms. Lausanne: Frontiers Media SA.
Kentner, Gerrit & Shravan Vasishth. 2016. Prosodic focus marking in silent reading: effects of discourse context and rhythm. Frontiers in Psychology 7. 319.
Köpcke, Klaus-Michael. 2002. Die sogenannte i-Derivation in der deutschen Gegenwartssprache: Ein Fall für outputorientierte Wortbildung. Zeitschrift für germanistische Linguistik 30. 293–309.
Kuijpers, Cecile & Wilma van Donselaar. 1998. The influence of rhythmic context on schwa epenthesis and schwa deletion in Dutch. Language and Speech 41. 87–108.
Kutscher, Anna. 2014. Intonation effects on adverb placement in German. In Proceedings of 'Exploring the Interfaces 3: Prosody and Constituent Structure'. Montreal: McGill University.
Leiner, Dominik J. 2014. Sosci Survey [Computer Software]. https://soscisurvey.de/ (accessed 12 January 2015).
Luo, Yingyi, Yunyan Duan & Xiaolin Zhou. 2015. Processing rhythmic pattern during Chinese sentence reading: An eye movement study. Frontiers in Psychology: Language Science 6. 1881.

Luo, Yingyi & Xiaolin Zhou. 2010. ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading. NeuroImage 49. 2836–2849.

Menninghaus, Winfried, Valentin Wagner, Eugen Wassiliwizky, Thomas Jacobsen & Christine A. Knoop. 2017. The emotional and aesthetic powers of parallelistic diction. Poetics. [article in press].

Myrberg, Sara. 2013. Sisterhood in prosodic branching. Phonology 30. 73–124.

Phillips, Nathaniel. 2017. YaRrr!: The Pirate's Guide to R. http://CRAN.R-project.org/package=yarrr (accessed 07 June 2017).

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://R-project.org (accessed 12 January 2016).

Rohdenburg, Günter. 2014. Grammatical variation and the avoidance of stress clashes in Northern Low German. In Jenny Cheshire & Dieter Stein (eds.), Taming the vernacular: From dialect to written standard language, 93–109. New York: Routledge.

Schlüter, Julia. 2005. Rhythmic grammar: the influence of rhythm on grammatical variation and change in English. Berlin: Mouton de Gruyter.

Selkirk, Elizabeth. 1984. Phonology and syntax. Cambridge, Mass.: MIT Press.

Selkirk, Elizabeth. 2000. The interactions of constraints on prosodic phrasing. In Merle Horne (ed.), Prosody: Theory and experiment, 231–261. Dordrecht: Kluwer Academic.

Shih, Stephanie, Jason Grafmiller, Richard Futrell & Joan Bresnan. 2015. Rhythm's role in genitive construction choice in spoken English. In Ralf Vogel & Ruben van de Vijver (eds.), Rhythm in Cognition and Grammar: A Germanic Perspective, 207–233. Berlin: De Gruyter.

Speyer, Augustin. 2010. Topicalization and stress clash avoidance in the history of English. Berlin: Mouton de Gruyter.

Vogel, Ralf, Ruben van de Vijver, Sonja Kotz, Anna Kutscher & Petra Wagner. 2015. Function words in rhythmic optimisation. In Ralf Vogel & Ruben van de Vijver (eds.), Rhythm in Cognition and Grammar: A Germanic Perspective, 255–276. Berlin: De Gruyter.

Wegener, Heide. 2004. Pizzas und Pizzen – die Pluralformen (un)assimilierter Fremdwörter im Deutschen. Zeitschrift für Sprachwissenschaft 23. 47–112.

Wiese, Richard. 2009. The grammar and typology of plural noun inflection in varieties of German. The Journal of Comparative Germanic Linguistics 12. 137–173.

Wiese, Richard. 2016. Prosodic parallelism – comparing spoken and written language. Frontiers in Psychology 7. 1598.

Wiese, Richard & Augustin Speyer. 2015. Prosodic parallelism explaining morphophonological variation in German. Linguistics 53. 525–559.

RHYTHMIC SEGMENTATION IN AUDITORY ILLUSIONS – EVIDENCE FROM CROSS-LINGUISTIC MONDEGREENS

Gerrit Kentner

Institut für Linguistik, Goethe-Universität Frankfurt
gerrit@lingua.uni-frankfurt.de

ABSTRACT

When confronted with non-native songs, listeners occasionally experience auditory illusions and perceive words in their native language although they are fully aware of the song lyrics being sung in a different language. We compiled two corpora with the original language of the song lyrics being English and the percept being either German or French. Against these two corpora, we tested the rhythmic segmentation hypothesis, specifically examining the cases of juncture misperceptions. The findings suggest that both German and French speakers use prominent syllables as anchors for segmentation, but they do so in language specific ways. For German listeners, prominent syllables signal the onset of lexical words. For French listeners, prominent syllables indicate phrase-finality. This cross-linguistic difference in boundary cueing corresponds with the specific role of prominent syllables in these languages, and makes a strong case for the concept of native listening in the context of sung speech.

Keywords: misperception, song lyrics, rhythmic segmentation, native listening

1. INTRODUCTION

Segmentation of the continuous speech stream into its component words is a task that listeners automatically perform not only when faced with spoken speech but also when confronted with sung speech.

The automaticity of segmentation in sung speech is vividly demonstrated by cross-linguistic mondegreens (misheard song lyrics) where listeners spontaneously perceive words or phrases in their native language even though they are fully aware that the song is actually being sung in a different language. This phenomenon has established itself as a popular meme and is a regular topic on radio shows where listeners report their misperceptions for comic effect. Given the automaticity of the auditory illusion [14], cross-linguistic mondegreens (also known as *soramimi*) can be considered an interesting test case for the study of speech perception in song (cf. [2]).

Previous work on speech segmentation suggests that listeners make use of heuristic segmentation procedures that are based on the experience with the structure of their native language [4]. One such heuristic is the metrical or rhythmic segmentation strategy (henceforth RSS) [3, 6]. According to the RSS, listeners use acoustically prominent (i.e. stressed or strong) syllables as anchors for speech segmentation but they do so in language-specific ways. Depending on the role of stress or prominence in the native prosodic system, prominent syllables may be more or less important for speech segmentation compared to other segmentation cues (e.g. phonotactics). Moreover, while in some languages, prominences indicate the beginnings of words or phrases, they may signal the end of comparable units in other languages.

In the case of German, a language with lexical stress and a preponderance of trochaic and dactylic lexical words in the native vocabulary [9], strong syllables are likely to be the initial syllables of lexical words. In contrast, weak or unstressed syllables typically represent grammatical words. Due to the strong propensity towards trochaic words (partly guaranteed by unstressed inflectional or derivational suffixes [16]), unstressed syllables that occur in lexical words are more likely to appear in non-initial positions (in spite of a comparatively rich system of unstressed prefixes).

French, in contrast, does not have clear lexical stress; rather, acoustically prominent syllables (typically those with a clear pitch excursion) have a demarcative function on the level of the phonological phrase rather than on the level of the word [12]. Phonological phrases roughly correspond to syntactic phrases (XPs) and obligatorily exhibit final prominence (‘primary phrase final accent’ [1], see also [12, 7]). A ‘secondary’ phrase-initial F0 rise is optional and not necessarily produced on the very first syllable of the phrase [12]. The initial F0-rise is a characteristic of, and possibly limited to, certain emphatic registers (e.g. political speeches) [8]. Thus, prominences should chiefly indicate phrase-finality in French.

Various studies attest the general validity of the

RSS across several languages using different research paradigms. Explicit evidence for the use of the RSS in the perception of song lyrics is, to the best of my knowledge, currently missing.

Most of the mondegreens we collected are more than one word in length. This offers the opportunity for misperceptions of word boundaries, so called juncture misperceptions [5]. We examine these juncture errors in two corpora of misheard song lyrics with the original language of the song being English and the percept being either German or French (henceforth English-German mondegreens and English-French mondegreens, respectively).

2. CORPUS STUDIES

2.1. Data and coding scheme

The data were gleaned from publicly available online platforms that archive radio programs in which listeners report misheard song lyrics. A considerable amount of these mondegreens is cross-linguistic. All mondegreens originate in pop songs of various genres that are accompanied by music. We were able to obtain 130 English-German mondegreens and 154 English-French mondegreens.[1] These were coded as follows. The mis-perceived part was aligned to the corresponding part of the original lyrics and the placement of word boundaries relative to the syllables was compared. Apart from the cases with coinciding word boundaries (36 English-German mondegreens, 51 English-French mondegreens), there are two types of possible boundary misperceptions: either a boundary insertion, i.e. a word boundary in the percept without corresponding word boundary in the original lyrics (1), or a boundary deletion, i.e. a single word in the percept spanning a word boundary in the corresponding original lyrics (2). Many mondegreens involve cases of both boundary insertions and deletions (3).

(1) Hope of de - li.ve - rance *orig. lyrics*
Hau auf die | Le.ber | wurst *Germ. percept*
"hit on the liver sausage"

(2) Let's live it | up *orig. lyrics*
Laisse les vé - los *Fr. percept*
"Leave the bikes"

(3) They | al - ways send the poor *orig. lyrics*
Thé - o | et ses deux porcs *Fr. percept*
"Theo and his two pigs"

The concept of word boundary used here is coarsly phonological: contractions to monosyllables, such as Engl. *you're* or Fr. *d'la* ∼ *de la* (Engl. '*of the*'), were treated as a single words. However, cliticized function words contributing their own syllable (engl: *kinda* ∼ *kind of*) were treated as separate words. Stems in compounds were counted as separate words, e.g. Germ. *Leber|wurst* (engl. '*liver sausage*'). Since the original lyrics and the percept are often not homophonous but rather distantly assonant, segmental detail, such as the affiliation of consonants to onsets or codas, was disregarded for the analysis.

Table 1 lists the number boundary misperceptions in the two corpora.

Table 1: Number of boundary misperceptions in the two corpora, broken down by type (insertion or deletion)

Percept	Insertions	Deletions	Total
German	50	87	137
French	75	99	174

In order to assess the validity of the RSS, the prominence of all syllables of the original (English) lyrics flanking the relevant word boundaries (inserted or deleted) were ascertained using an online dictionary of American English [13]. A syllable was marked as strong or stressed if it is quoted as having primary or secondary stress within the word it appears in. All other syllables were marked as weak or unstressed. Note that establishing the syllable prominence this way disregards its actual prominence in the song. Since, however, there does not seem to be a trivial way to assess the actual prominence of syllables in a song, the present approach seems to be an adequate approximation (any metric for actual syllable prominence in songs would certainly need to take into account, apart from lexical stress, absolute and relative acoustic measures on and around that syllable, let alone the characteristics of the instrumentation; complicating matters even more, the relative contribution of the various factors may be variable and song-specific).

2.2. Predictions

The general prediction of the RSS is that the distribution of word boundaries in inter-lingual auditory illusions based on non-native (here: English) song is dependent on the distribution of prominences in the lyrics. The specific dependency of boundary distribution and the distribution of prominences is subject to the role of prosodic prominence in the native linguistic system and thus differs between language groups.

2.2.1. *English-German mondegreens*

German is a language with lexical stress, i.e. basic prominences are assigned at the lexical level. Content words in German exhibit a strong tendency towards a trochaic pattern [9, 16]. Compared to German, the native vocabulary of English has a stronger tendency for lexical words to be stressed monosyllables [10]. Correspondingly, German listeners experiencing auditory illusions on the basis of English song lyrics are predicted to segment the sung input before strong syllables and these strong syllables then represent the initial syllables of lexical words in the percept. Moreover, German listeners are predicted to integrate unstressed syllables with preceding strong syllables into trochees; therefore, before unstressed syllables, boundaries that are present in the original lyrics are likely to be deleted.

As in English, grammatical words in German tend to be monosyllabic and unstressed. If boundaries are inserted before weak syllables, these weak syllables are likely to represent grammatical words.

2.2.2. *English-French mondegreens*

The French prosodic system assigns prominences at the phrasal level with the final syllable of every phrase carrying a boundary tone. The RSS would predict French listeners converting English song lyrics into French illusions to use the lexical prominences of the original English lyrics primarily as signal to a final boundary. That is, boundary insertions should preferably be placed after stressed syllables. Unstressed syllables, in contrast, are typically not phrase-final in French. Correspondingly, word boundaries after unstressed syllables in the original lyrics should be deleted in the French illusion. French prosodic phrases roughly correspond to syntactic phrases and syntactic phrases are likely to start in grammatical words (e.g. noun phrases, prepositional phrases, verb phrases in the periphrastic past tense (*passé composé*) typically start in function words: Det N, Prep Det N, Aux V). The English-French mondegreens should reflect this. If taken as phrase-final boundary tones, the prosodic prominences are likely to be followed by a (phrase-initial) grammatical word.

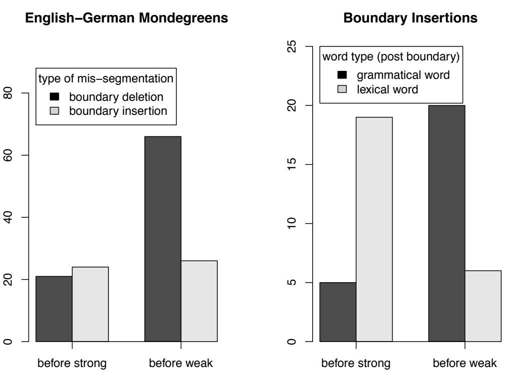
2.3. Results

2.3.1. *English-German mondegreens*

Boundary insertions occur nearly equally before weak syllables and before strong syllables, while boundary deletions are clearly more common, especially before weak syllables as compared to strong syllables, cf. Figure 1, left panel. A linear mixed model with binomial link function was applied to test the influence of following syllable prominence (strong vs weak) on the type of mis-segmentation (boundary deletion vs insertion); to account for the fact that cases of mis-segmentation from the same mondegreen are not independent, the song and the specific mondegreen were included as random factors. This model yields a significant main effect of syllable prominence (*β=-1.072, SE=.387, z=-2.77, p=.0058*), confirming the prediction that German listeners tend to insert boundaries before strong syllables and especially avoid word boundaries before weak syllables.

The right panel of Figure 1 shows all cases of boundary insertions broken down by the prominence of the following syllable. The plot suggests a clear dependence of word class and syllable prominence. That is, as predicted, boundaries inserted before strong syllables tend to produce lexical words while boundaries inserted before weak syllables produce grammatical words. A second linear mixed model probing whether the prominence of the syllable following a mis-inserted boundary reliably predicts the word class, confirms this dependence (*β=2.539, SE=0.685, z=3.706, p=.0002*).

Figure 1: Left panel: Mis-segmentations in English-German mondegreens broken down by type (deletion / insertion) and prominence of following syllable. Right panel: Boundary insertions broken down by type of following word (lexical / grammatical) and prominence of following syllable.

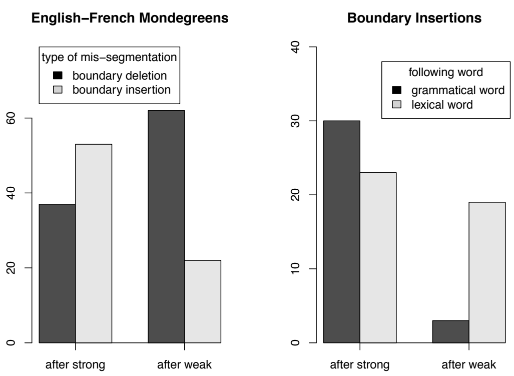


2.3.2. *English-French mondegreens*

As in the English-German mondegreens, there are more boundary deletions than insertions in the English-French mondegreens. Again, the distribution of boundary insertions and deletions is dependent on the position of strong syllables (cf. Figure

2, left panel). When compared to the original English lyrics, boundaries in the French percept are preferably inserted after strong syllables and deleted after weak syllables. A mixed logistic regression with song and mondegreen included as random effects confirms the significance of this interaction (*β=-1.396, SE=0.328, z=-4.257, p<.001*). The right panel of Figure 2 shows the boundary insertions broken down by the prominence of the preceding syllable and the word type of the following word. All in all, post-boundary words are more likely to be lexical words. However, in accordance with the predictions, when the inserted boundary is placed after a strong syllable, the following word is clearly more likely to be a grammatical word. The interaction between pre-boundary syllable prominence and post-boundary word class is highly significant (*β=-2.112, SE=0.68, z=-3.104, p=0.0019*).

Figure 2: Left panel: Mis-segmentations in English-French mondegreens broken down by type (deletion / insertion) and prominence of preceding syllable. Right panel: Boundary insertions broken down by type of following word (lexical / grammatical) and prominence of preceding syllable.



3. DISCUSSION

The analyses of the English-German and English-French mondegreens provide a strong case for the general validity of the RSS in non-native song perception. In both corpora, clear dependencies between the type of juncture misperception (insertion or deletion) and the prominence of the surrounding syllables were found. That is, both German and French native speakers use strong syllables as anchors for segmentation when confronted with English songs. Crucially however, as predicted, the specific anchoring of strong syllables differs according to the role of strong syllables in the native prosodic system. For German listeners, strong syllables signal onsets of lexical words, in accordance with the preponderance of the trochaic foot structure in the German vocabulary. For the same reason, word boundaries before weak syllables are clearly avoided, as attested by the high proportion of boundary deletions in this context. Most likely, the overall bias for boundary deletions reflects the general tendency for lexical words in German to be trochaic compared to the preponderance of monosyllables in English [10]. Importantly, again in accordance with the RSS, the distribution of boundary insertions relative to surrounding syllable prominence predicts the word class of the post-boundary word, i.e. boundaries before strong syllables produce lexical words and boundaries before weak syllables produce grammatical words.

In contrast to German listeners, French listeners tend to posit boundaries not before but after strong syllables. Given the obligatory phrase-final accent of French, strong syllable are most likely not taken to correspond to lexical stress but to phrase-final prominence. Therefore, boundaries after strong syllables chiefly correspond to phrase boundaries, and only by implication to junctures between words. Since phrases preferably start in grammatical words, a grammatical word is predicted to follow a boundary inserted after a strong syllable – and this tendency is indeed confirmed in the mondegreen data.

Finally, a cautionary note about the nature of the data is in order. Since the mondegreens originate from songs accompanied with music, the misperceptions may well be based on the structure of the ambient music. Research on song perception has shown that linguistic and musical perception are strongly intertwined [11]. In general, strong positions in the music (on-beat notes) likely coincide with strong positions in the original lyrics (on-beat syllables) but this correlation is certainly not perfect. The present study does not consider this potential nuisance factor, so it remains unclear to what extent the juncture misperceptions originate from the text or from the tune (cf. [15] for a similar point). Moreover, it has to be noted that this natural experiment lacks the control over the materials that would be desirable. The English-German and English-French mondegreens stem from different songs and are therefore, strictly speaking, incomparable; however, given that the results neatly confirm the predictions derived from the prosodic systems of the two languages, it is highly unlikely that the differences found are specifically due to the different English source songs. In all probability, the mondegreens reflect the true difference between the prosodic systems of German and French and thus confirm the use of the native linguistic system in non-native song perception.

4. REFERENCES

[1] Astésano, C., Bard, E. G., Turk, A. 2007. Structural influences on initial accent placement in French. *Language and Speech* 50, 423–446.

[2] Beck, C., Kardatzki, B., Ethofer, T. 2014 Mondegreens and soramimi as a method to induce misperceptions of speech content – Influence of familiarity, wittiness, and language competence. *PloS One* 9.1, e84667.

[3] Cutler, A., 1990. Exploiting prosodic probabilities in speech segmentation. In: Altmann, G. T. M. (ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge: MIT Press, 105–121.

[4] Cutler, A. 2012. *Native listening: Language experience and the recognition of spoken words*. Cambridge: MIT Press.

[5] Cutler, A., Butterfield, S. 1992. Rhyhtmic cues to speech segmentation: Evidence from juncture misperceptions. *J. Mem. Lang.* 31, 218–236.

[6] Cutler, A., Norris, D. 1988. The role of strong syllables in segmentation for lexical access. *J. Exp. Psy: HPP* 14, 113–121.

[7] Di Christo, A., 1998. Intonation in French. In: Hirst, D., Di Christo, A. (eds.), *Intonation systems: A survey of twenty languages*. Cambridge University Press, 195–218.

[8] Di Christo, A. 2000. Vers une modélisation de l'accentuation en français. Deuxième partie: Le modèle. *Journal of French Language Studies*, 10, 27–44.

[9] Eisenberg, P. 1991. Syllabische Struktur und Wortakzent. *Zeitschrift für Sprachwissenschaft*, 10, 37–64.

[10] Fucks, W. 1955. *Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen.* Wiesbaden: Springer VS.

[11] Gordon, R. L., Schön, D., Magne, C., Astésano, C., Besson, M. 2010. Words and melody are intertwined in perception of sung words: EEG and behavioral evidence. *PLoS One*, 5.3, e9889.

[12] Jun, S. A., Fougeron, C. 2000. A Phonological model of French intonation. In: Botinis, A. (ed.), *Intonation: Analysis, Modeling and Technology*. Dordrecht : Kluwer, 209–242.

[13] Merriam Webster online dictionary. http://www.merriam-webster.com/dictionary/. (retrieved January 2015.)

[14] Otake, T. 2007. Interlingual near homophonic words and phrases in L2 listening: Evidence from misheard song lyrics. Proc. 16th ICPhS. Saarbrücken. 777–780.

[15] Vitevitch, M. S. 2002. Naturalistic and experimental analyses of word frequency and neighborhood density effects in slips of the ear. *Language and Speech*, 45.4, 407-434.

[16] Wiese, R. 1996. *The phonology of German*. Oxford: Clarendon Press.

[1] I thank Kypriani Sinaris who collected and coded the English-French mondegreens.
Both corpora will be made available on request.