Poetic Rhyme Reflects Cross-Linguistic Differences in Information Structure[☆]

Michael Wagner[a,*], Katherine McCurdy[b]

[a] *McGill University*
[b] *Harvard University*

Abstract

Identical rhymes (right/write) are considered satisfactory and even artistic in French poetry but are considered unsatisfactory in English. This has been a consistent generalization over the course of centuries, a surprising fact given that other aspects of poetic form in French were happily applied in English. This paper puts forward the hypothesis that this difference is not merely one of poetic tradition, but is grounded in the distinct ways in which information-structure affects prosody in the two languages. A study of rhyme usage in poetry and a perception experiment confirm that native speakers' intuitions about rhyming in the two languages indeed differ, and a further perception experiment supports the hypothesis that this fact is due to a constraint on prosody that is active in English but not in French. The findings suggest that certain forms of artistic expression in poetry are influenced, and even constrained, by more general properties of a language.

Keywords: rhyme, information structure, focus, givenness, poetry

[☆]This research was supported by FQRSC Grant NP-132516: La prosodie: production, perception et différences interlinguistiques; and a Canada Research Chair in Speech and Language Processing (Grant 212482). Thanks to Jonathan Abramsohn, T.C. Chen, Lizzie Smith, Steffanie Scheer, and Jozina vander Klok for help with conducting the experiment and data analysis. We would like to thank Luc Baronian, Lev Blumenfeld, and Bob Ladd for their very insightful and detailed feedback on an earlier version of this paper. Thanks also for helpful comments and suggestions to Marc Brunelle, Mara Breen, Marie-Hélêne Côté, Anne Cutler, John Bowers, Meghan Clayards, Stephen B. Cushman, Judith Degen, Heather Goad, Florian Jaeger, Paul Kiparsky, Aniruddh Patel, Donca Steriade, and Brian Trehearne, to the audiences at talks at Chicago University, Concordia, Carleton University, Princeton, Radboud University, UMass Amherst, and the Max Planck Institute for Psycholinguistics in Nijmegen, and to three very helpful reviewers.

[*] *Corresponding Address:* Michael Wagner, McGill Linguistics, 1085 Dr. Penfield Avenue, Montréal, QC H3A 1A7, chael@mcgill.ca

1. Rhymes, Identical and Other

Rhymes can be defined as a pair of words that are phonologically identical from the last accented vowel to the end of a word (light/night); they typically occur at the end of a line in poetry (Fabb, 1997, 118). Identity rhyme—a rhyme in which the syllable onsets preceding the accented vowels are identical (write/right, attire/retire)—is commonly used in French poetry, while in English poetry it is considered to be "unconventional and even unacceptable" (Small, 1990, 141) and to "fall ridiculously flat" (Hollander, 1985, 118).

Poetic devices such as rhyme and alliteration (words beginning with the same onsets) have been argued to not just enhance aesthetic experience but also to affect comprehension and recall (Lea et al., 2008). Allopenna et al. (1998) found that rhyming competitors are activated in word recognition, suggesting that rhyme plays a role in the organization of the mental lexicon. Steriade (2008) presents evidence that rhymes are relevant for the phonology of a language even outside of poetry. None of these extra-poetic functions of rhyme, however, have been shown to explain the cross-linguistic differences between what counts as a good rhyme.

Hollander (1989, 14) employs an instance of a rhyme consisting of two homophonous words—a special case of an identity rhyme—in order to advise against its usage:

(1) The weakest way in which two words can chime
Is with the most expected kind of rhyme—
(If it's the only rhyme that you can write,

A homophone will never sound quite right.)

Holtman (1996, 187) and Small (1990) argue that the scarce uses of identity rhyme attested in English generally reveal an awareness that they violate an expectation. This is similar to a conscious violation of a metrical expectation in order to convey a poetic effect, which is sometimes seen in poetry with a fixed meter (Halle and Keyser, 1971). In other words, both the scarcity and the nature of use of identity rhyme in English poetry reflect its stigmatized status. An antipathy for identity rhyme in English may have existed as early as 1584, when King James issued a treatise proscribing the practice: "That ye ryme nocht twyse in ane syllabe. As for exemple, that ye make not prove and reprove rhyme together, nor hove for hoveing on hors bak, and behove." (see Rait 1900 for the original text).

Identical rhymes have to be distinguished from repetitions of the same word, since repetition obeys quite different regularities (and has different poetic effects) from rhyme (**?**). We will consider only identical rhymes of words that differ in meaning.

Interestingly, it is only identical rhymes, i.e., those rhymes preceded by identical onsets (right/write, called 'rimes trs riches' in **?**) that are considered weak, while rhymes that merely extend into the onsets but do not have identical onsets ('rimes riches:' train/rain) are unexceptional and quite commonly used in English:

(2) I have looked down the saddest city lane.
[...]
And dropped my eyes, unwilling to explain.

From: Robert Frost, Acquainted with the Night

In French, in contrast to English, identical rhymes are unexceptional and often said to be even superior to simple rhymes. Aroui (2005) notes that identical rhymes do not seem to be used for a particular effect or with a particular pattern of recurrence, suggesting they are considered normal rhymes.[1] It is easy to find identical rhymes in French poetry, for example they occur quite frequently in the poetry of Émile Nelligan, a poet from Québec:

(3) [...]
vocalise d'une voix d'eau d'or
[...]
Soupire et rit dans la nuit qui dort.
From: Émile Nelligan, Vasque

The first part of this paper aims to establish that indeed the languages differ in their rhyming repertoire, first by looking at the usage of identical rhymes in English and French, and second by using experimental evidence that native speakers of the two languages sharply differ in their intuitions about the quality of identical rhymes. This difference is surprising given the persistent influence of French poetry on English poetry. The second part of this paper proposes a novel account that relates the difference in identity rhyme usage

[1] Repetitions, on the other hand, are considered a banal form of rhyme also in French (Elwert, 1965, 88). According to Elwert identical rhymes that are morphologically related are also considered weaker by some.

to a difference in how prosody reflects information structure in the two languages (Ladd, 2008), and presents supporting experimental evidence for this explanation.

2. French and English Poets Differ in their Use of Rhyme: A Natural Experiment

How different are the usages of rhymes in English and French? Since the poetry produced by individual authors varies along many dimensions, it is not easy to assess whether and to what extent these two languages differ in their overall use of rhymes, especially since modern poetry often does not employ rhyme at all. In order to quantify the difference in a more controlled way, we looked at translations of a German children's book, Wilhelm Busch's *Max & Moritz* (first published 1865), which comprises 208 couplets, all of which rhyme and none of which are identical rhymes. In German, identical rhymes are considered weak, just like in English.

We chose this particular book because we assumed that the genre of a humorous (albeit a bit gruesome) children's book would allow for a playful use of rhymes, so we expected substantial variation in rhyme usage across different translators. Also, we were confident that there would be a sufficient number of translations into both languages to compare the variability of rhyme usage within a language against the variability across language boundaries. The corpus of translations of this book constitutes a natural experiment in the usage of different rhyme types.

2.1. Materials and Methods

We were able to obtain 6 translations into English and 5 into French (listed in the appendix). Almost all translations were rhymed and consisted of a comparable number of couplets. One French translation was very loose and used hardly any rhymes, so we excluded it from analysis. The other books were scanned, and the text was hand-annotated for rhyme types by the authors and double-checked by a research assistant.

2.2. Results & Discussion

The distribution of rhyme in our mini-corpus confirm that there is a dramatic difference in the usage of identical rhymes between the two languages. Table 1 summarizes the usage of rhymes in different translations. In English, many translations have no identical rhymes, like the German original, one had 1 (0.5%) and another 3 (1%). In French, on the other hand, identical rhymes account for 16–36% of all couplets.

This consistent difference in identity rhyme usage between all English and French translators contrasts with the usage of 'rimes riches' in the same translations. Rimes riches are used with comparable frequency across all three languages (an average of 3.5% of the rhymes in the English translations and 2.8% in French, compared to 3.4% in the original), while poets within languages vary quite a bit in their use (e.g., between 1.9% and 7% in English).

Given the small and unequal sample size and possible difference in variance, we used Welch's t-test (independent, two-tailed, two-sample) in order to test for significance. The average proportion of identical rhymes in English *vs.* French were significantly different ($t(df \approx 3.01) = -4.8, p < 0.02$).

Table 1: *Rhyme Usage by Language*

	German	English Translations						French Translations			
		I	II	III	IV	V	VI	I	II	III	IV
total	208	208	198	211	207	188	205	208	209	174	202
rimes riches (%)	3.4	3	5.6	1.9	3.9	2.1	4.3	3.4	2.0	2.0	3.9
identity (%)	0	0.5	0	0	0	0	1.4	35.5	29.2	16.3	15.5

The difference in proportions of rimes riches, however, (on average there were slightly *more* in English) was not significant ($t(df \approx 7.9) = 0.85, p < 0.42$).

The analysis of our mini-corpus of translations confirms that there is a dramatic difference in rhyme usage between English and French in that identical rhymes are avoided in English but are used quite frequently in French; however, rimes riches are used with comparable frequency.

2.3. The Role of the Lexical Rhyming Resources of a Language

When assessing the rhyme inventories of a language it is very informative to consider the lexical statistics and phonotactics. In a language like French, in which word-stress is always final, a rhyme always involves the final part of the last syllable of a line starting from the stressed vowel: a 'masculine rhyme.' However, in a language such as English in which stress can fall on pre-final syllables, this is just a special case; rhymes more generally include all material from the last accented vowel to the end of the line, and feminine rhymes (e.g., 'double rhymes' like blended/mended, or 'triple rimes' like cereal/material).[2] In addition, French has much more restricted phonotactics,

[2]In cases in which the last accent does not fall on the last word, a rhyme can even include multiple words, a phenomenon often called 'mosaic rhyme.' Here's one from a Max & Moritz translation:

so the number of possible rhymes overall is substantially smaller.

Given the clear differences in their phonology, could it be that identical rhymes are stigmatized in English because they are simply exceedingly rare compared to the case of French? Maybe rhyme like a *pear/pair* are bad because there are not enough identical alternatives to choose from, as Luc Baronian (p.c.) and a reviewer suggested. Explanations based on lexical resources were used in Hanson and Kiparsky (1996) to explain how languages pick a particular poetic meter, and it seems plausible that rhyming patterns might work similarly. (Hanson and Kiparsky, 1996) argue that there is a balance between the fit between lexicon and meter (language select meters in which their lexical resources are usable in the greatest variety of ways) and interest (all-too obvious poetic tools are not aesthetic).

In order to check whether there is a simple explanation for the status of identical rhymes in French and English we estimated the likelihood of rhymes based on word corpora. The French lexicon in Lexique (New et al., 2004) of 142,693 words partitions into 624 rhyme cohorts with a median length of 9, and 4,077 identity rhyme cohorts, with a median length of 4. The English lexicon of 160,595 word forms in Celex (Baayen et al., 1995) partitions into 40,903 rhyme cohorts with a median length of 1, and 62,681

(i) Hence, the village folk commend him
And are eager to befriend him.

For an interesting discussion of this type of rhyme see Hook (2008), with further cross-linguistic evidence that rhymes must be defined based on the location of the last accentual peak, just like in English and French.

identity rhyme cohorts, also with a median length of 1. Clearly, the languages differ dramatically in their rhyming resources, but an obvious explanation for why identity rhyme in particular should be stigmatized in English does not emerge from these numbers: If rhymes in language were good when they are likely to occur by accident, then English should not be a rhyming language at all, since rhymes are hard to come by and they are comparatively contrived; if rhymes were better when they were *infrequent* because they're harder to find and hence more aesthetic, then identical rhymes should be *better* than non-identical rhymes, because they're harder to find in both French and English.

Most identical rhymes in the French translations are non-homophonous identical rhymes. In English, even non-homophonous identical rhymes are considered weaker than normal rhymes; for example, many speakers find *moat/remote*, *retire/attire*, and *saloon/balloon* to be weak rhymes, although these pairings may not be as bad as fully homophonous identical rhymes.[3] In our mini-corpus 3 out of 4 French translations had homophonous rhymes (2 on average) while only one one out 6 English translation had any homophonous rhymes.[4]

[3]It might also be that *remote/moat* is worse than *retire/attire* because only one word contains a distinguishing additional syllable, as a reviewer pointed out. As we will see, our experiments included only one non-homophonous rhyme in each language. For these, we did not find a difference, but more data would be necessary here. See also footnote 8 on French.

[4]It contained the same rhyme twice: *two/too*. This is an interesting rhyme because the two words occur in syntactically very different positions. The word 'two' was part of an NP argument, while 'too' attaches at the sentence level. This difference results

Given the low number of uses, could it be that homophone-rhymes are avoided also in French? In order to estimate whether the usage of homophonous identical rhymes in French was higher or smaller than expected by chance, we estimated how likely it is that an identity rhyme is a homophonous identity rhyme. We found that about 0.01% of the identity rhymes in Lexique are homophonous rhymes, while in our mini-corpus of poetry translations an average 4.9% of identical rhymes were homophonous, suggesting that homophonous identical rhymes are used much more frequently than expected based on their probability, and it seems thus that they are not avoided in French.

The question of how exactly to quantify rhyme likelihood is complex. For example, one might want to consider word frequencies, morphological relatedness and other factors.[5] A thorough analysis would easily fill a separate article on the topic. However, it seems safe to conclude that French and English differ in their use of identity rhyme, and that the lexical and phonological differences alone do not provide an obvious explanation for this difference.

in a substantial acoustic difference in terms of length and pitch, which makes them less identical, and hence less of an identity rhyme.

[5]See also an insightful blog-post by Mark Liberman on the Language log: http://languagelog.ldc.upenn.edu/nll/?p=1946

3. French and English Native Speakers Differ in their Intuitions about Rhyme

identical rhymes are all but absent in English, and it is generally assumed that this is not because they are scarce but because they are deemed poor and are therefore avoided. In order to establish whether identical rhymes are indeed considered unsatisfactory by English speakers and satisfactory by French speakers, a rating experiment was conducted in which participants listened to and evaluated recorded couplets containing three different rhyme types.

3.1. Participants

Three groups participated in the experiments: native speakers of North American English (born and raised in Canada or the US), native speakers of Québec French, and native speakers of European French. Each group consisted of 24 participants, except for English, where we accidentally ran 25 participants. We excluded three European French speakers and five Qubec French speakers because they were born or spent part of their childhood somewhere other than France or Qubec respectively. We included both European and Québec French speakers in this experiment because we thought that greater exposure to English might exert an influence on Québec French speakers—we will return to this point later. Most participants were run in the phonetics lab at McGill University, but due to difficulties in recruiting French-speaking participants we ran 14 of our Québec speakers and eight European French speakers in a public building in Montréal, and 12 of our European French participants were run in a public library in Aix-en-Provence,

France.

3.2. Materials and Procedure.

Each participant listened to 15 mini-poems. The items varied by three conditions across participants:[6]

(4) *Identical Rhyme:*
The gardener watered the soil, then rose
and picked a single crimson rose.
Good Rhyme:
Pat inhaled deeply through her nose
and picked a single crimson rose.
Bad Rhyme
She strolled through the garden when she woke
and picked a single crimson rose.

All stimuli were original compositions. In both English and French, all identical rhymes but one were homophonous, but all differed in their meaning. We focused on homophonous identical rhymes because they form a particularly spectacular illustrations of the difference between the two languages. We tried to avoid identical rhymes that were similar in meaning since semantic resonance might interact with the quality of a rhyme (see Wimsatt, 1954). The English stimuli were recorded by a female native speaker of English, and the French stimuli by a female native speaker of European French.

[6]The sound stimuli and a list of all the items are posted at http://prosodylab.org/∼chael/papers/rhyme/.

Participants were told that the rhymes were chosen by non-native speakers, and they were to evaluate whether these rhymes were satisfactory rhymes in English/French based on their native-speaker intuitions (they were debriefed after the experiment). This was intended to put participants into a position of feeling like an 'expert' qualified to evaluate the rhymes.

Each experiment commenced with a practice session of four couplets, to familiarize participants with the procedure. Participants listened to each stimulus via Logitech USB headset, and evaluated the acceptability of the rhyme on a scale ranging from 1 (very bad) to 5 (very good) by clicking the appropriate numbered box on the screen. The experiment was run using experimental scripts in the speech analysis program Praat (Boersma and Weenink, 1996).

3.3. Results

The plots displayed in Figure 1 show that English speakers rate identical rhymes as being relatively unacceptable, while both Québec French speakers and European French speakers do not reliably distinguish in acceptability between identical rhymes and good rhymes.

The data were analyzed using a mixed-model regression analysis, controlling for item and subject as random effects, and adding condition ('good', 'identity', 'bad'), language ('English', 'EurFrench', 'QueFrench') and their interaction as fixed effects.[7] We tested the significance of the interaction

[7]We used the 'lmer' function of the lme4 package in R. The model we used looked as follows: model.lm < − lmer(response ∼ language*condition + (1|item) + (1|subject)). Baayen et al. (2008) note that in a mixed-model regression a comparison can be considered

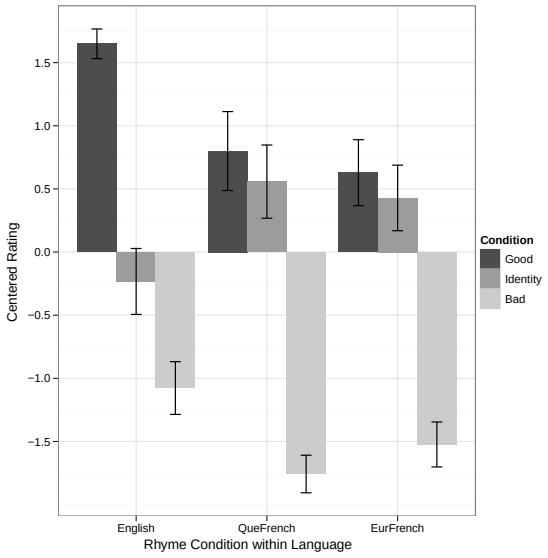

Figure 1: Average centered ratings obtained in the rhyme experiments. Participants rated the utterances on a scale from 1 (very poor) to 5 (very good); the ratings were centered for analysis to a scale ranging from -2 to 2.

between condition and language by comparing a regression model including the interaction and one excluding it using log-likelihood ratio test, showing a highly significant difference ($\chi^2(4) = 93.7, p < 0.001$). The difference in mean rating between 'good' and 'identity' in English differs significantly from the difference in rating in these two conditions in European French ($t = 5.4, p < 0.001$) and Québec French ($t = 5.0, p < 0.001$). We also computed mixed models within each language, and the difference between 'good' and 'identity' was significant in English ($t = -14.13, p < 0.001$), but not in European French ($t = -1.4, p < 0.15$) or Québec French ($t = -1.7, p < 0.09$).

The results are just as expected given the hypothesis—but could it be that factors other than phonological identity influence the judgments? In our English data, 6 out of the 15 identical rhymes involved morphologically related words, and one of the unrelated rhymes was orthographically identical. One might think that morphologically related rhymes are worse than less related identical rhymes. However, there was no significant difference between the two groups of items in English: morphologically unrelated and orthographically distinct identical rhymes were rated just as bad as morphologically related or orthographically indistinct ones, suggesting that the infelicity of identical rhymes is not driven by morphological or orthographic factors.[8]

significant if the t-value for a comparison exceeds the absolute value 2. In addition, we also report a conservative estimate of the p-value based on mcmc-sampling, using the pvals.fnc function of the languageR R-package.

[8]Thanks to Marie-Hélène Côté for pointing out that two of our rhymes rhyme in European French but not in Québec French pronunciation. Since our speaker was European French, it is unlikely that this would have affected the outcome. The response pattern for

Only one item each in English and French included a non-homophonous identity rhyme. These items showed the same pattern—the non-homophonous identity rhyme was rated as much worse than the good control in English but rated as good (in fact, even slightly better than the good control) in French. This suggests that non-homophonous rhymes pattern no differently from homophonous ones: identical rhymes are bad in English but good in French.[9]

3.4. Discussion

Our findings confirm the widely held assumption that identity rhyme is a satisfactory form of rhyme to native speakers of French, but not to native speakers of English. Although not significantly different from the European French pattern, the pattern of the Québec French speakers tends a bit more in the direction of English—we will return to this difference below.[10]

---

the Québec listeners did not show any sign that they treated them differently than the European listeners.

[9]In order to further test whether homophonous *vs.* non-homophonous identical rhymes are different, we recorded a set of 17 French couplets with identical rhymes from our corpus; 6 were homophonous, 5 rhymed by virtue of a grammatical ending, and 6 were other non-homophonous identical rhymes. We had them rated by 12 native speakers of French (6 from Québec and 6 from France). The mean ratings were between 4.0 and 4.3 for the three groups, with no significant difference between them, suggesting that all three types of identical rhymes are considered good in French. We have not yet conducted a comparable study for English.

[10]A reviewer points out that the particular meter of our poems may have added to the observed effects, since a prominent beat at the end of the line adds salience to them. A follow-up manipulating different meters could test this idea. It seems unlikely, however, that the effect would completely disappear with a different meter.

4. A Prosodic Difference Between English and French

Why did identity rhymes not catch on in English, despite the substantial and sustained influence of French poetry on English writing over the course of centuries? The influence was so strong—and asymmetric—that Ezra Pound quipped in a 1913 article that "the history of English poetic glory is a history of successful steals from the French" (cited after Pondrom 1974). Does the difference in opinion about identical rhymes reflect mere aesthetic or stylistic variation in poetic traditions, or even, as Richardson (1909) argues, the force of King James's very decree upon English literary practice?

Our hypothesis is that differences in information-structural effects on prosody are the actual explanation of this difference. In English and French, the last accent in an utterance usually falls on the stressed syllable of the last word. In English, however, words or constituents that are highly accessible in the discourse (or "given") often remain unaccented, or are "destressed," and have reduced prominence (cf. Halliday 1967; Selkirk 1995; Schwarzschild 1999). See Cutler (1997) and Wagner and Watson (2010), for overviews of the experimental literature on the topic, and and Xu and Xu (2005) for a recent discussion of the phonetic realization of this type of reduction. In fact, destressing given material is usually obligatory when it is possible. This can be seen as a result of the 'given-new contract' (Clark and Haviland, 1977), which requires that when it is possible to mark information as given and link it to an antecedent in the discourse context it must be marked as such (cf. Williams, 1997). We will refer to this phenomenon henceforth as *anaphoric destressing* (following Rooth, 1996, i.a.):

(5) An AMERICAN farmer met a CANADIAN farmer.

Accents are marked with small-caps, destressing with underlining. While the last accent within each noun phrase would usually fall upon the last word of the phrase (American FARMER, Canadian FARMER), in (5) the word farmer remains unaccented, highlighting the informational contrast between *American* and *Canadian* (cf. Ladd, 2008, and references therein). This type of anaphoric destressing, however, has been shown not to occur in various Romance languages (Ladd, 2008; Swerts et al., 2002; Swerts, 2007), including French. In example (6), accentual prominence remains on the rightmost content word in both phrases (*américain*), even though this information is contextually given.[11].

(6) Un flic AMÉRICAIN a rencontré un fermier AMÉRICAIN.
a policeman American has met a farmer American
'An American policeman met an American farmer.'

Information structure thus does not affect prosody in French the same way as it does in English. But how does that relate to identical rhymes? Identical rhymes differ from typical cases of anaphoric destressing in that they are odd even when they differ in meaning, as in (1). If anaphoric destressing prohibits accents on constituents encoding contextually given information, why would this be relevant for words that merely sound the same but mean something

[11]Note that in a French noun phrase the adjective usually follows the noun it modifies, in contrast to English, but this is not crucial here. See (Ladd, 2008) for discussion of a variety of examples with parallel word orders.

different, and even for words that are identical only from the accented syllable on?

This brings us to an interesting quirk of English (and other Germanic languages): focus/givenness-marking seems to have been generalized to givenness at the phonological-form-level. Ladd (2008, 234), for example, gives the following observation from a BBC broadcast, in which stress on *Titanic* shifts to the first syllable marking the contrast to *Brittanic*:

(7) Greek divers have found the wreck of the British liner Brittanic, sister ship of the TItanic ...

Williams (1980, 1997) observes, crucially, that there are cases in which a sentence is infelicitous when two adjacent expressions end with an accented word that is phonologically identical.[12] In (8c), semantically, an accent on the final word should be acceptable because it contrasts with another word in the context (just as in (8a) and (8b)), but the fact that the previous clause happens to end with the same accented word prohibits that pronunciation:

(8) a. JOHN hugged MARY, and then MARY hugged JOHN.
b. JOHN hugged MARY, and then JOHN was hugged by HER.
c. #JOHN hugged MARY, and then JOHN was hugged by MARY.

The utterance sounds odd because 'Mary' is not deaccented, just as if it was given information that is accented–but in fact it is semantically contrastive.

[12]The effect seems to be strongest the antecedent was at the end of a bigger previous prosodic domain.

The infelicity ensues because it is 'phonologically given.' We refer to this odd phenomenon as the 'Williams-Effect.' If the purpose of prosodic backgrounding in English is to mark what semantic information is given—as is usually assumed—then this effect constitutes a 'bug' of this system, and constitutes a givenness-illusion.

Our claim is that the infelicity of identical rhymes is due to the Williams-Effect. In other words, identical rhymes are actually fine rhymes, but couplets ending with an identical rhyme sound poor because they violate the prosodic constraints of English which require given information to be deaccented, including phonologically given information. If this is correct, then only languages that show the Williams-effect should show a stigmatization of identity rhyme.

The Williams-effect is likely to be orthogonal to the 'repeated names penalty' observed in the literature on the usage of pronouns versus full proper names (Gordon et al., 1993; Gordon and Chan, 1995). The use of a full name as opposed to a pronoun has been shown to result in longer reading times, both in subject and direct object position, when the previous sentences had a co-referent subject. Based on this characterization of the effect, all sentences in the paradigm in (8) should incur a repeated-names penalty because *John* is repeated. Also, it would be unclear why deaccenting *by Mary* substantially improves (8c).[13] And furthermore, this alone would not explain why, at least

[13]In a production study, not reported here, we found that speakers pronounce sentences like (8a) and (8b) with an accent on the the final word, while in (8c) prominence shifts to the preposition *by* or the predicate *hugged*, so the infelicity of (8c) is indeed at least to a large extent due to a lack of anaphoric destressing.

according to Williams, the effect can also be observed with pronouns:

(9) a. JOHN hugged MARY, and then MARY hugged JOHN.
b. JOHN hugged MARY, and then JOHN was hugged by HER.
c. #JOHN hugged HER, and then JOHN was hugged by HER.

Let's suppose nevertheless that it was indeed the case that the paradigm in (8) illustrates purely an effect of the repeated-names penalty—then English and French should not differ with respect to the Williams-effect, since French has also been reported to show a repeated-names penalty (Fossard 1999). Our hypothesis makes a different prediction: if identical rhymes indeed are good in French and bad in English, then the Williams-effect should exist in English but not in French. The following section reports a perception experiment testing for the Williams-effect in both English and French.

5. The Willliams-effect and the (In)Felicity of identical rhymes

Our second perception experiment tested for the presence of the Williams-effect in non-poetic contexts in all three languages. Based on our hypothesis that identical rhymes in English are considered weak because of the Williams-effect, we predicted that it should be present in English, just as Williams (1980) hypothesized, and absent or at least weaker in French.

5.1. Materials and Method

Our stimuli consisted of two sentences conjoined by *and*. Again, there were three conditions: this time 'contrast,' 'Williams,' and 'anaphoric.' In the 'contrast' condition, an accented final noun phrase contrasted with the

noun phrase ending the previous clause. In our 'Williams' condition, both sentences ended with the same accented NP, the second instance contrasting with the NP carrying the same thematic role in the previous sentence. Finally, in our 'anaphoric' condition, both ended with the same accented NP, without any contrast:

(10) *Contrast:* JOHN hit SUE, and then JOHN was hit by MARY.
Williams: JOHN hit MARY, and then JOHN was hit by MARY.
Anaphoric: JOHN saw MARY, and then JOHN was hit by MARY.

The accent on *Mary* in the 'contrast' condition was as expected, since Mary encodes new information (and it may be employed in contrast to *Sue* or *John* here). In the 'anaphoric' condition the accent should be infelicitous: since *Mary* encodes old information, the name should be deaccented. In our 'Williams' condition the contrast to *John* should in principle license the accent on the second instance of *Mary*, despite the fact that it encodes discourse-salient information, just like in condition 'contrast.' However, we predicted that English speakers would find it infelicitous due to the Williams-effect.

If French indeed lacks anaphoric detressing, then a different pattern is expected. French speakers were predicted to rate both the 'Williams' condition and the 'anaphoric' condition as more acceptable compared to English speakers. Any deprecation of these conditions would have to be purely due to a repeated-names penalty, a much weaker effect than a failure to do anaphoric detressing does in English.

In both English and French, the experiment consisted of 24 items in the respective languages, varying by three conditions. Stimuli were recorded by the same speakers as in the rhyme experiment and were rated by the same listeners as in the rhyme experiment. In order to avoid participants guessing that this experiment was somehow related to rhyme, we ran it before the rhyme experiment with each subject.[14] Participants again listened to recordings via headsets and evaluated the acceptability of each stimulus on a scale from 1 to 5, using an experimental script in Praat.

5.2. Results and Discussion

Figure 2 illustrates the results, which confirmed our predictions, with some qualifications. It is not surprising that even the sentences in the 'anaphoric' condition were not rated as very bad (very bad would have been -2, but the mean is around 0.03), since according to our hypothesis this condition involved an odd pronunciation of an otherwise acceptable sentence. It is quite clear, however, that the sentence in condition 'anaphoric' were rated worse on average than the sentences in condition 'contrast.'

A mixed-model analysis including condition, language, and their interaction as fixed effects, and subject and item as random effects, showed a clear interaction between condition and language. The interaction was highly significant based on a log-likelihood-comparison between a model including the interaction and one excluding the interaction ($\chi^2(4) = 51.1, p < 0.001$). More

[14] In order to assure that the order of experiments did not influence the responses we ran an additional group of 12 English native speakers only on the rhyme experiment, replicating the results of experiment 1.

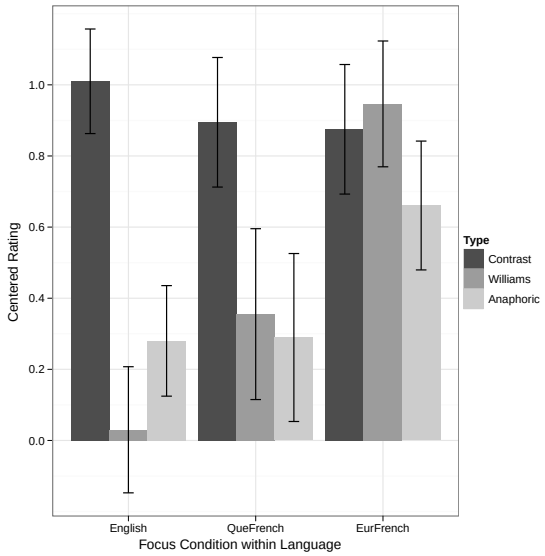

Figure 2: Average centered ratings obtained in the focus experiment. Participants rated the utterances on a scale from 1 (very poor) to 5 (very good); the ratings were centered for analysis from -2 (very poor) to 2 (very good). As the mean values for our centered data all fell between 0 and 1, we display only this range, for ease of comparison.

specifically, the difference between 'contrast' and 'Williams' in English differs significantly from the difference between these two conditions in European French ($t = 7.2, p < 0.004$) and Québec French ($t = 2.9, p < 0.001$). These results are according our predictions.

When looking within language, we found that in English, as expected, 'Williams' is significantly worse than 'contrast' ($t = -9.8, p < 0.001$), but not so in European French ($t = 0.65, p < 0.52$). Unexpectedly, however, Québec French patterns with English here in showing a significant difference for this comparison ($t = -4.7, p < 0.001$). Similarly, our 'anaphoric' condition turned out to be considered quite bad in Québec French ($t = -5.2, p <$

0.001).[15] While the differences between 'Williams' and the other two conditions are significantly smaller compared to English, it seems as if Québec French came out half-way between the English and the European French pattern.

One possible explanation for this difference between Québec French and European French is that the former group has had more exposure to English and may therefore be influenced by the use of anaphoric destressing in that language. In our language questionnaire, speakers of Québec French reported higher proficiency in English compared to European French speakers. However, it could also be that Québec French simply differs from European French in the way prosody is affected by information structure.[16]

We further tested our hypothesis by looking at the correlation between the degree to which there is a Williams-effect and the degree to which identical rhymes are considered bad for individual subjects. We computed the mean of the z-score of the ratings per condition for each subject in the two experiments. Then we tested how well the mean ratings for the 'Williams' condition in the focus experiment and the 'identity' condition in the rhyme experiment correlated. As predicted, the correlation between the two mea-

[15] It approached significance in European French ($t = -1.9, p < 0.06$) as well. We interpret this as an effect of the repeated-names penalty—note, however, that the means in European French and Québec French are closer than in English, and significantly so according to the mixed model.

[16] A reviewer pointed out that it would be interesting to test how rhyme intuitions change depending on L2 proficiency. This could be of interest both in their native language and in the target language. There are a number of other factors that could be looked at, for example age of exposure might be relevant as well.

sures is significant, with $R^2 = 0.13; F(1, 63) = 9.8; p < 0.003$. Given the small n for this analysis (the data from every participant is reduced to one data point), it is quite striking that we found a significant correlation nevertheless. The correlation was strongest in the Québec French group ($R^2 = 0.17$). In other words: a particular subject's rating of identical rhymes correlated a subject's rating of Williams-sentences, suggesting that the two phenomena are closely related to each other.

Even though there was a correlation between the two effects, the Williams-effect in Québec French was stronger than the weak antipathy for identity rhyme would lead one to expect. Maybe this is due to the fact that Québec French speakers get a lot of positive evidence that identity rhyme is deemed acceptable in French—as we saw, identical rhymes are very common in French poetry. For example, a children's song well-known in Québec rhymes *dents* 'teeth' with *dedans* 'within.' So those Québecois speakers that show a Williams-effect may rate identical rhymes as better than would otherwise be expected because they have learned by experience that they are deemed good rhymes.

6. General Discussion

Despite centuries of sustained mutual influence between French and English poetry, identity rhyme remains very common within one poetic tradition and marginalized in the other. That the two languages indeed differ dramatically in poetic practice in this regard was confirmed by looking at a set of translations of the same children's book, a natural experiment in rhyme usage. We then presented evidence from an experiment showing that identity rhyme is

deemed satisfactory by native speakers of French but not by native speakers of English.

Our proposed explanation for the difference is that identical rhymes sound odd in English because of the overgeneralization of anaphoric destressing first pointed out by Williams. Our second experiment provided the first experimental confirmation of the Williams-effect in English, and also showed that it is absent in European French, and much less pronounced in Québec French. At an individual level, there is a correlation between the degree to which native speakers show a Williams-effect and the degree to which they reject identical rhymes.

If our hypothesis is correct, we would expect other Germanic languages to pattern with English, since they show similar patterns with respect to how prosody is affected by information structure, and other Romance languages to pattern with French. While we have not explored these cross-linguistic predictions, suggestive evidence comes from the *Max & Moritz* mini-corpus. For example, the original German text indeed contains no identical rhymes.[17]

[17]Grimm (1850) notes that identical rhyme or 'rhrender Reim' was used in Middle-High German, although Paul (1893, 114) holds that it was frowned upon already then. A fair number of the rhyme examples Grimm discusses have identical final syllables but do not have final stress, and thus wouldn't count as identical rhyme in the narrow sense. According to our hypothesis, if Grimm is right and identical rhyme was acceptable, this constitutes evidence that anaphoric destressing must not have been active in German yet; however, if Paul is correct it must already have been part of the grammar. Rhyme usage might thus reveal something about aspects of pronunciation that are hard to diagnose based on written sources otherwise. There are at least 10 translations of Max & Moritz into German dialects, and one could test our predictions more based on these.

Spanish, on the other hand, patterns with French in its lack of anaphoric destressing (Ortiz-Lira, 1995), so we would predict identity rhyme to be permissible. And indeed, two Spanish translations that we annotated contain 12% and 12.2% of identical rhymes respectively. While this rate of identity rhyme usage may be smaller than typical values in French, it is more than six times higher than the rate observed in any Germanic version of this poem. Given that Spanish does not always have final stress like French, identical rhymes are much less common in the lexicon, so 12% is a substantial proportion.[18] More cross-linguistic data could further test our claim that the acceptability of identical rhymes correlates with prosodic focus-effects.

This paper argues for an intrinsic link between prosodic information-structure effects and constraints on rhymes. We did not offer an explanation of why anaphoric destressing should exist in Germanic languages but not in French and other Romance languages—this is a question that needs to be explored independently. A number of differences between English and French might be relevant here, since they may well influence the use of the prosodic effects of focus and givenness and/or rhyme. The intriguing expectation based on the results of this study is that whatever will explain the difference in the first will by implication account for the second.

One possibility is that English and French differ both in their information structure and in their rhyme inventory because of their different prosodic systems. In French, the accent (almost) always falls on the last syllable of

[18]None of the identical rhymes in Spanish were homophones, although many involved a single word rhyme (like *remote/moat* in English. Homophones are rare in Spanish compared to French, so one cannot conclude from this that homophonic rhymes are avoided.

a sentence, and the phonology of the language revolves around accentual phrases rather than domains of word stress as in English (Jun and Fougeron, 2000). However, it cannot be the particulars of French phonology alone that explain its lack of anaphoric destressing, since Italian and Spanish have word-stress systems but both lack anaphoric destressing. This also speaks against an explanation of a lack of destressing in terms of a 'destressing-deafness,' as it was reported for French in Dupoux et al. (1997). French native speakers were found to ignore differences in accent placement, in contrast to Spanish speakers who were found to be sensitive to stress location—but if this were to explain the lack of anaphoric destressing, then Spanish should pattern with English in this regard, contrary to fact. For the same reason an explanation in terms of the likelihood of homophones seems doubtful. While French is has a high number of homophones compared to English—a well-known problem for automatic speech recognition in French (see Lamel and Gauvain, 1993)—other Romance languages seem to pattern with English in terms of the likelihood of speech recognition errors resulting from homophony (Barnett et al., 1996), so homophone frequency does not appear to correlate with the presence/absence of anaphoric detressing.

A possible reason Romance languages might work differently in their prosodic information structuring is that they are highly inflected and word-stress tends to fall on one of the last syllables. This has the effect that sentences ending with identical rhymes by virtue of their grammatical endings (so called 'homoeoteleutons') occur with some frequency. So maybe applying an English-style focus constraint in a Romance language would result in too many 'false alarms' due to the Williams-effect, that is, deaccenting for

phonological reasons would be quite frequent rather than being the exception as in English. This hypothesis seems quite plausible, and would provide an explanation in terms of lexical resources after all, but one in terms of how they interact with information structuring rather than in terms of how they directly influence the likelihood of certain rhymes. In fact, Wimsatt (1954) relates the fact that Chaucer employed identity rhyme quite frequently to the fact that Middle English still had more stressed suffixal endings (see also: Holtman 1996: 177). This type of explanation would only explain the absence of anaphoric destressing, however, if somehow an English-style anaphoric destressing rule *necessarily* goes hand-in-hand with the Williams-effect, which current theories of focus-marking would not lead one to expect.[19]

The contribution of this paper is to show that an otherwise puzzling difference in the rhyming patterns in French and English can be explained as an effect of an independently established difference in anaphoric destressing—the question of what explains this difference in information structure itself remains open. That patterns of artistic expressions are grounded in linguistic patterns of the artist's native language has also been found in music (Patel and Daniele, 2003), and it should come as no surprise then if the same holds true for linguistically expressed art. The restrictions on identical rhymes across languages constitute further evidence that a better understanding of

[19] An additional factor that could be relevant is that syllable structure is delineated more crisply in French (Cutler et al., 1986), and plays a crucial role in speech segmentation. A difference in segmentation strategies could affect intuitions about rhymes, although it is not obvious how this will translate into an alternative explanation for the patterns observed here, or the correlation with information structuring.

the linguistic system of a language can illuminate the study of poetry and vice-versa, as advocated by Jakobson (1960), and that "a good number of what we think of as traditional and arbitrary conventions [on poetic form] are anchored in grammatical form, and seem to be, at the bottom, a consequence of how language itself is structured" (Kiparsky, 1973, 11).

Appendix: Max & Moritz and its Translations

German Original: Wilhelm Busch, 1865. Max & Moritz. Eine Bubengeschichte in sieben Streichen. Reprinted in: Max Görlach (Ed.), 1994: Max & Moritz polyglott. 12th edition (first edition 1982). München: Deutscher Taschenbuchverlag.

English 1: Walter W. Arndt, 1982. Max & Moritz. A story of two rascals in seven tricks. In: W. W. A., The Genius of Wilhelm Busch. The Regents of California Press. Reprinted in: Görlach, 1994.

English 2: Elly Miller, 1981. Mac and Murray. A Tale of Two Rascals, in Seven Episodes. Reprinted in: Görlach (Ed.), 1986: Wilhelm Busch's Max and Moritz in English Dialects and Creoles. Hamburg: Buske.

English 3: Charles T. Brooks, 1871. Max and Maurice. A Juvenile History in Seven Tricks. New York: Roberts.

English 4: Wilhelm Busch, 2003. Max and Moritz and Other Bad-Boy Stories and Tricks. Translated from the German by Andy Gaus. Rockville, MD: James A. Rock & Co.

English 5: Wilhelm Busch, 1962. Max and Moritz. With many more mischief makers more or less human or approximately animal. Edited, annotated, and translated by H. Arthur Klein and others. New York: Dover.

English 6: Wilhelm Busch, 1996. Max und Moritz auf englisch. Englische Nachdichtung von Percy Reynolds (Max and Moritz. A Tale of Two Scamps in Seven Pranks). Stuttgart: Reclam.

French 1: Jean Amsler, 1981. Max et Maurice. Histoire de gamements en sept farces. First publication in: Görlach, 1994.

French 2: Henri Mertz, 1982. Max et Maurice. Histoire de deux petits espi'egles. In: Görlach, M., 1994b: Max und Moritz in Romanischen Sprachen. Essen, Blaue Eule.

French 3: Wilhelm Busch, 1978. Max et Moritz. Adapté de l'allemand par Cavanna. Paris: Mouche.

French 4: André Thérive, 1952. Max et Maurice, ou les sept mauvais tours de deux petits garçons. Adapté par A. T. Paris: Ernst Flammarion. Reprint: Munich, Braun and Schneider, 1965.

French 5: Duchesne, Christiane, 2002. Max et Maurice en sept mauvais coups. Adapté librement de Wilhelm Busch.

Spanish 1: Víctor Canicio, 1982: Max y Moritz. Una historieta en siete travesuras. In: Görlach, M., 1982.

Spanish 2: Rosa Enciso und Guido Mensching, 1990. Paco y Pedro. La historia de dos pillos es siete travesuras traducida por R.E. y G.M. In: Görlach, M., 1994b.

More translations of Max & Moritz into these languages are listed in Görlach (1994), but we have not yet been able to obtain them.

References

Allopenna, P., Magnuson, J., Tanenhaus, M., 1998. Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. Journal of Memory and Language 38, 419–439.

Aroui, J.L., 2005. Rime et richesse des rimes en versification française classique, in: Murat, M., Dangel, J. (Eds.), Poétique de la rime.. Champion, Paris.

Baayen, R., Davidson, D., Bates, D., 2008. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language 59, 390–412.

Baayen, R.H., Piepenbrock, R., Gulikers, L., 1995. The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Barnett, J., Corrada, A., Gao, G., Gillick, L., Ito, Y., Lowe, S., Manganaro, L., Peskin, B., 1996. Multilingual speech recognition at Dragon Systems, in: Proceedings of the Fourth International Conference on Spoken Language Processing.

Boersma, P., Weenink, D., 1996. PRAAT, a system for doing phonetics by computer. report 132. Institute of Phonetic Sciences of the University of Amsterdam.

636 Clark, H.H., Haviland, S., 1977. Comprehension and the given-new con-
637 tract, in: Freedle, R.O. (Ed.), Discourse production and comprehension.
638 Erlbaum, Hillsdale, pp. 1–40.

639 Cutler, A., 1997. The comparative perspective on spoken-language process-
640 ing. Speech Communication 21, 3–15.

641 Cutler, A., Mehler, J., Norris, D., Segui, J., 1986. The syllable's differing
642 role in the segmentation of French and English* 1. Journal of memory and
643 language 25, 385–400.

644 Dupoux, E., Pallier, C., Sebastian, N., Mehler, J., 1997. A destressing 'deaf-
645 ness' in French. Journal of Memory and Language 36.

646 Elwert, W.T., 1965. Traité de versification française des origines à nos jours.
647 Klincksieck.

648 Fabb, N., 1997. Linguistics and literature: language in the verbal arts of the
649 world. Blackwell, London.

650 Gordon, P.C., Chan, D., 1995. Pronouns, passives, and discourse coherence.
651 Journal of Memory and Language 34, 216–231.

652 Gordon, P.C., Grosz, B.J., Gilliom, L.A., 1993. Pronouns, names, and the
653 centering of attention in discourse. Cognitive science 17, 311–347.

654 Halle, M., Keyser, S.J., 1971. English stress: Its form, its growth, and its
655 role in verse. Harper & Row, New York.

656 Halliday, M., 1967. Notes on transitivity and theme in English, part 2.
657 Journal of Linguistics 3, 199–244.

Hanson, K., Kiparsky, P., 1996. A parametric theory of poetic meter. Language , 287–335.

Hollander, J., 1985. Vision and Resonance. Oxford University Press. 2nd edition.

Hollander, J., 1989. Rhyme's reason: A guide to English verse. Yale University Press.

Holtman, A.I., 1996. A generative theory of rhyme: A optimality approach. Ph.D. thesis. Universiteit Utrecht.

Hook, P.E., 2008. Basic word order, information peak, and the use of rhyme in traditional prosodies, in: Proceedings of the International Symposium: Language & Literature and Language & Power, July 9 – 11, 2007, Akita University. pp. 1–6.

Jakobson, R., 1960. Linguistics and poetics. Style in language , 350–377.

Jun, S., Fougeron, C., 2000. A phonological model of French intonation. Intonation: Analysis, modelling and technology , 209–242.

Kiparsky, P., 1973. The role of linguistics in a theory of poetry. Daedalus 102, 231–244.

Ladd, D.R., 2008. Intonational Phonology. Cambridge University Press. 2nd edition.

Lamel, L., Gauvain, J., 1993. Cross-lingual experiments with phone recognition, in: Proceedings of IEEE ICASSP-93, Citeseer.

679 Lea, R., Rapp, D., Elfenbein, A., Mitchel, A., Romine, R., 2008. Sweet
680 Silent Thought: Alliteration and Resonance in Poetry Comprehension.
681 Psychological Science 19, 709–716.

682 New, B., Pallier, C., Brysbaert, M., Ferrand, L., 2004. Lexique 2: A new
683 french lexical database. Behavior Research Methods, Instruments, & Com-
684 puters 36.

685 Ortiz-Lira, H., 1995. Nucleus placement in English and Spanish: A pilot
686 study, in: Lewis, J.W. (Ed.), Studies in General and English Phonetics.
687 Essays in honor of J.D. O'Connor. Routledge, London & New York, pp.
688 255–256.

689 Patel, A., Daniele, J., 2003. An empirical comparison of rhythm in language
690 and music. Cognition 87, 35–45.

691 Paul, H., 1893. Grundriss der Germanischen Philologie. Trübner.

692 Pondrom, C.N., 1974. The road from Paris; French influence on English
693 poetry, 1900-1920. Cambridge University Press.

694 Rait, R.S. (Ed.), 1900. A Royal Rhetorician: A treatise on Scottis Poesie. A
695 Counterblaste to Tobacco. etc. etc. A. Constable and Co., Westminter.

696 Richardson, C.F., 1909. A Study of English Rhyme. Dartmouth College,
697 Hanover, NH.

698 Rooth, M., 1996. Focus, in: Lappin, S. (Ed.), Handbook of Semantics.
699 Blackwell, London, pp. 271–297.

Schwarzschild, R., 1999. Givenness, AVOIDF and other constraints on the placement of accent. Natural Language Semantics 7, 141–177.

Selkirk, E.O., 1995. Sentence prosody: Intonation, stress, and phrasing, in: Goldsmith, J.A. (Ed.), Handbook of Phonological Theory. Blackwell, London, pp. 550–569.

Small, J.J., 1990. Positive as Sound: Emily Dickinson's rhyme. The University of Georgia Press, Athens, GA.

Steriade, D., 2008. Light and heavy clusters: weight and total duration. Paper presented at the Clusters & Complexity Workshop, IPS, Munich.

Swerts, M., 2007. Contrast and accent in Dutch and Romanian. Journal of Phonetics 35, 380–397.

Swerts, M., Krahmer, E., Avesani, C., 2002. Prosodic marking of information status in Dutch and Italian: A comparative analysis. Journal of Phonetics 30, 629–654.

Wagner, M., Watson, D.G., 2010. Experimental and theoretical advances in prosody: A review. Language and Cognitive Processes (online version available since May 26) .

Williams, E., 1980. Remarks on stress and anaphora. Journal of Linguistic Research 1, 1–16.

Williams, E., 1997. Blocking and anaphora. Linguistic Inquiry 28, 577–628.

Wimsatt, W.K., 1954. On the relation of rhyme to reason, in: Beardsley, M., Wimsatt, W. (Eds.), The Verbal Icon: Studies in the Meaning of Poetry. University of Kentucky Press, Lexington, pp. 152–166.

Xu, Y., Xu, C., 2005. Phonetic realization of focus in English declarative intonation. Journal of Phonetics 33, 159–197.