

英文题目: Stress, information, and language typology*

中文题目: 重音, 信息, 和语言的分类

中文原载: 《语言科学》第6卷第5期(总第30期)3-16, 2007年9月

英文翻译者: 端木三

Author's address:

San Duanmu

Department of Linguistics

University of Michigan

Ann Arbor, MI 48109

USA

中文摘要:

重音分词重音和语句重音。前者可因语言而异。后者是否因语言而异, 看法不一。本文先介绍几个语句重音理论, 并指出其弱点, 然后提出一个新的理论, 称为“信息-重音原则”。根据新的理论, 语句重音由信息量决定: 信息多的词要重读, 信息少的词不用重读。词的信息量, 可以由对话人和对话环境决定, 也可以由信息理论, 句法结构, 和上下文决定。新的理论, 既比前人的简单, 又能克服其弱点, 还能简化语言的分类。

Abstract:

Stress theory covers word stress and phrasal stress. The former can differ from language to language. Whether the latter can also differ from language to language is still an open question. This article reviews several theories of phrasal stress and points out their shortcomings. Then a new theory is proposed, which is called the 'Information-Stress Principle', according to which phrasal stress is determined by information load: words with more information should be stressed and words with less information need not. The information load of a word can be determined by the speaker and his/her environment; it can also be determined by information theory, syntactic structure, and the context. The new theory is simpler than previous ones, without their shortcomings, and offers a simpler view of language typology.

中文关键词:

信息-重音原则, 重音, 信息论, 信息量, 语言参数, 语言分类

Key words:

Information-Stress Principle, stress, information theory, information load, language parameters, language typology

1. Introduction

Every language has stress. In some languages, stress can distinguish word meanings, as shown in (1). In some languages, stress is not used to distinguish word meanings, but it still occurs, such as in contrastive stress, exemplified in (2). For clarity I use underline to indicate stressed syllables.

- (1) Stress used to distinguish word meanings (English)

Noun: *content*

Verb: *content*

- (2) Contrastive stress (Shanghai Chinese)

wo bu xing Wang, wo xing Huang

I not named Wang, I named Huang

‘My name is not Wang, but Huang.’

A stressed syllable is usually longer, with a full rime (no reduction), and often a certain pitch contour (Fry 1955, 1958), or wider pitch range (Shen 1985, Liu and Xu 2005). However, in tone languages (such as Chinese and Japanese), since pitch is used to contrast word meanings, native judgment for stress is often unclear.

Most stress theories often focus on word stress (e.g. Wang and Feng 2006). Less discussion is given to stress above the word level. In this study I focus on the latter, which I shall refer to as phrasal stress, which covers both compound stress and phrasal stress.

2. Theories of phrasal stress

In this section I review main theories of phrasal stress in the past 40 years.

2.1. Chomsky and Halle (1968): Two phrasal stress rules

Chomsky and Halle (1968) discuss English and propose two rules, which are the Compound Stress Rule and the Nuclear Stress Rule, re-phrased in (3) and (4), where A and B are the immediate constituents of a compound or phrase, and underline indicates words that receive stress from the rules.

- (3) Compound Stress Rule:

In a compound [A B], if B is a compound, then B gets stress, otherwise A gets stressed.

Examples:

[<u>A</u> B]	<i><u>black</u>-board</i>
[A [<u>B</u> C]]	<i>stone <u>black</u>-board</i>
[[<u>A</u> B] C]	<i><u>black</u>-board store</i>

- (4) Nuclear Stress Rule:

In [A B], B gets stress.

Examples:

[A <u>B</u>]	<i>many <u>people</u></i>
[A [B <u>C</u>]]	<i>many old <u>people</u></i>
[[A <u>B</u>] C]	<i>many <u>people</u> <u>arrived</u></i>

Chomsky and Halle (1968) did not discuss stress in other languages, nor did they explain why compounds and phrases have different stress rules. For them, therefore, it is possible that in some languages compounds and phrases both have left-headed stress (i.e. stress on the left), or both have right-headed stress (i.e. stress on the right), or compounds have right-headed stress and phrases have left-headed stress.

2.2. Halle and Vergnaud (1987): Stress parameters

Chomsky (1981) proposes that differences among languages can be attributed to different settings of a few parameters. Following this idea, Halle and Vergnaud (1987) propose that stress differences in different languages can be attributed to a set of stress parameters. Two such parameters are shown in (5).

- (5) Stress parameters (Halle and Vergnaud 1987):
 [+/-bounded] whether the length of a foot is limited
 [left/right] whether stress in a foot is on the left or on the right

For example, we can translate the rule in (4) as the parameter settings in (6).

- (6) Nuclear Stress Rule:
 [-bounded, right] unlimited foot length, stress on the right

It can be seen that the parameter settings of Halle and Vergnaud (1987) do not depend on syntax. For example, when choosing stress parameters for compounds, some languages can choose left-headed stress, and some can choose right-headed stress. Therefore, in principle some languages can have stress patterns that are opposite to those in English: compounds are right-headed (instead of left-headed), and phrases are left-headed (instead of right-headed). However, the focus of Halle and Vergnaud (1987) is still word stress. Although they assume that phrasal stress can differ in different languages, no examples are given.

2.3. Duanmu (1990): Non-head Stress

Duanmu (1990) propose that the two phrasal stress rules of Chomsky and Halle (1968) can be combined into one, shown in (7), where underline indicates a constituent that receives stress.

- (7) Non-head Stress:
 In a syntactic structure [A B], one is the head and the other the non-head. Phrasal stress is assigned to the non-head.
 Examples: [X YP]
 [YP X]
 [ZP [X YP]]

In most cases, the Non-head Stress rule gives similar results as the two rules of Chomsky and Halle (1968), as the examples in (8) show.

- (8) Illustration of the Non-head Stress rule
 Compounds:
 [A B] *black-board*
 [A [B C]] *stone black-board* (not [A [B C]])
 [[A B] C] *black-board store*

Phrases:

[P NP] *in school*
 [V NP] *watch movies*
 [AP (F) NP] *pretty (F) woman* (not [AP N])

For compounds, there is a slight difference for the structure [A [B C]], where the result of Duanmu seems to agree with native intuition better. In phrases, the structure [AP (F) NP] follows the analysis of current syntactic theory (e.g. Pollock 1989, Ritter 1991, and Cinque 1993), where the syntactic head is not the noun, as traditionally assumed, but a functional element F (similar to the *de* in Chinese). Additional discussion of the role of Non-head Stress in Chinese can be found in Duanmu (1999b).

2.4. Cinque (1993): Cumulative stress

Cinque (1993) proposes that phrasal stress is directly related to the depth of syntactic structure. Stress starts from the deepest syntactic units (the innermost brackets) and accumulates along the way. An example is shown in (9), where stress is assigned in three cyclic steps, starting from the innermost word, and the taller the column of * marks, the more stress a word has.

(9) Cumulative stress

Cycle 1	Cycle 2	Cycle 3
		*
	*	(. *)
*	(. *)	(. (. *))
(*)	(. (*))	(. (. (*)))
[* [* [*]]]	[* [* [*]]]	[* [* [*]]]

The proposal of Cinque (1993) is very similar to the Non-head Stress rule of Duanmu (1990), because in a syntactic structure, the syntactic non-head always has a deeper structure than the syntactic head. This can be seen in (10).

(10)

/\	/\
X YP	YP X
/\	/\
...	...

The syntactic non-head has a deeper structure because it is phrasal in nature and can expand. In contrast, the syntactic head is a word (or a morpheme), which cannot expand.

2.5. Ladd (1996): Phrasal stress can differ in different languages

According to Duanmu (1990) and Cinque (1993), phrasal stress can be derived from syntax, and it works the same way for all languages. Ladd (1996: 174) offers a different view. He believes that ‘accentuation ... is a matter of the grammar of specific languages rather than of universal principles’. Ladd does not discuss in which ways phrasal stress can differ from language to language, but offers some hypothetical examples, shown in (11).

(11) Hypothetical examples of variation in phrasal stress (uppercase indicates stress)

Language A	Language B
<i>This is book RED</i>	<i>This is BOOK red</i>
<i>I bought car NEW</i>	<i>I bought CAR new</i>
<i>He has nose BIG</i>	<i>He has NOSE big</i>

The purpose of (11) is to show that even if two languages have exactly the same syntax (in terms of word order), they can still differ in phrasal stress. Ladd also offers some other examples, to be discussed later.

2.6. Zubizarreta and Vergnaud (2000): Focus stress

Like Duanmu (1990) and Cinque (1993), Zubizarreta and Vergnaud (2000) believe that phrasal stress can be derived from syntax. However, unlike previous people, who treat normal stress and focus separately, Zubizarreta and Vergnaud (2000) propose that normal stress and focus can be treated in the same way.

Zubizarreta and Vergnaud (2000) adopts Chomsky's (1971) definition of focus, which refers to 'new information'. It can be shown by answers to question sentences. A focus can be large or small, ranging from one word, as in (12a), to an entire sentence, as in (12b).

(12) Focus (new information, indicated by underline)

- a. Question: Who watched a movie?
 Answer: John watched a movie.
- b. Question: What's the news today?
 Answer: John bought a book.

In (12a), the focus of the answer is *John*. In (12b), the focus of the answer is the entire sentence. The two answers have the same syntax, but the main stress falls on different words: it is on *John* in (12a) and *book* in (12b).

When the focus is the entire sentence, the stress rule of Zubizarreta and Vergnaud (2000) is given in (13), which is similar to those of Duanmu (1990) and Cinque (1993).

(13) Assign main prominence to the first selectional dependent.

Example: [XP X], [X XP], [XP [XP X]]

Or: Assign main prominence to the category that is lowest in the asymmetric c-command ordering.

Example: [XP X], [X XP]

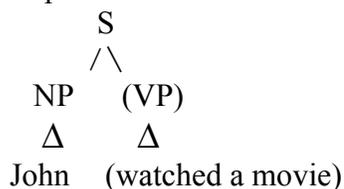
The 'selected dependent' and 'c-commanded category' are similar to the syntactic non-head of Duanmu (1990), and the 'lowest category' is similar to the deepest constituent of Cinque (1993). In the answer of (12b), *book* is the lowest category, which is also the location of main stress. Therefore, the prediction of stress assignment is correct.

When the focus is smaller than the entire sentence, how is stress assigned? Zubizarreta and Vergnaud (2000) propose a de-accent rule, rephrased in (14).

- (14) The de-accent rule:
Anaphoric material is de-accented.

In (12a), *watched a movie* is the anaphoric material, which is de-accented. In the tree representation in (15), the de-accented part is shown in parentheses.

- (15) Representation of de-accenting



The original stress inside the parentheses will be ignored. As a result, only *John* is visible, and therefore it carries main stress.

3. Problems with current theories

There are some common problems in the theories just reviewed. First, why should some constituents (such as the syntactic non-head) have more stress than others? No answer is offered. Second, why should there be more stress on the focus? There is no explicit answer either. Third, it has been noted that frequent words often have less stress than infrequent words. For example, Fidelholtz (1975) offers the examples in (16), where the sounds of interest are underlined.

- (16) Frequent: *information* [ɔ̃] astronomy [ə]
 Infrequent: *importation* [oɔ̃] gastronomy [æ]

The word pair *information* and *importation* are similar in structure, yet there is stress reduction in the former but not in the latter. Similarly, *astronomy* and *gastronomy* are also similar in structure, yet there is stress reduction in the former but not in the latter. Why should frequent word have less stress than infrequent ones? There is no explicit explanation in stress theories either.

4. The Information-Stress Principle

I propose a new theory of phrasal stress, which I call the Information-Stress Principle. It offers a uniformed explanation of the problems just mentioned. Let us consider contrastive stress first, exemplified in (17), where underline indicates words under contrast.

- (17) Contrastive stress:
 a. *My name is Huang, not Wang.*
 b. *John didn't buy a movie; he borrowed one.*

In (17a) contrastive stress falls on *Huang* and *Wang*. In (17b) contrastive stress falls on *buy* and *borrowed*. Words under contrast are spoken with greater prominence and are intuitively quite clear to native speakers. One would assume that all languages have contrastive stress, and that all contrastive stress works the same way, namely, words under contrast have more stress than other words. In particular, we do not expect any language that uses an opposite version of contrastive stress (anti-contrastive stress), as shown in (18), where underline indicates stressed words.

- (18) Anti-contrastive stress (hypothetical case)
My name is Huang, not Wang.

In (18), the words under contrast, *Huang* and *Wang*, are spoken with no stress, whereas other words are spoken with stress. I am not aware of any such language.

Why, then, do all languages use the stress pattern in (17) and no language uses the stress pattern in (18)? The answer is quite obvious: words under contrast have more information of interest, and more stress should be given to words that carry more information. Although this principle seems quite simple, it can explain a lot of problems. Duanmu (2005) calls it the Information-Stress Principle, which is defined in (19).

- (19) The Information-Stress Principle:
 Words with more information are spoken with more stress.

In (17), the words under contrast (and hence with more information) are spoken with more stress, which satisfies the Information-Stress Principle. In (18) the words under contrast are not spoken with more stress, which violates the Information-Stress Principle; therefore, (18) is unnatural.

Next we consider how information load is determined. According to Information Theory (Shannon 1948), the information load of a sign depends on the probability of its occurrence. Therefore, we can use (20) to define information load.

- (20) Information load:
 a. The more likely a word is found, the less information load it has.
 b. The more possible words there are for a given syntactic position, the less probability each word has, and the more information load it has for that position.

The definition of information load in effect says that the more unexpected something is, the more information it provides. This concept is intuitively natural. Consider the news events in (21).

- (21) a. The Chinese table tennis team won the championship.
 b. The Chinese soccer team won the championship!
 c. A patient woke up after seven years in a coma!!!

The news in (21a) is highly expected, and therefore it provides little new information. The news in (21b) is unexpected, because its probability is much lower than that of (21a). Therefore the information load of (21b) is high. The news in (21c) is extremely unexpected, and therefore it carries even more information load.

Let us take a closer look at how information load is calculated. Consider the English sentence in (22), where we focus on the information load of the last two words.

- (22)
- | | | |
|-----------------------|------------|-----------|
| | (He has) a | car. |
| Word category: | article | noun |
| Number of choices: | 2 | thousands |
| Probability per word: | high | low |
| Information load: | low | high |

A singular noun in English can only take one of two articles (either *a* or *the*), but there are thousands of nouns. Therefore, in the article position, the probability of each article is very high, and its information load is very low. In the noun position, on the other hand, the probability for each noun is very low, and so its information load is very high. Similarly, consider the Chinese sentence in (23), where we focus on the last two words.

(23)	(ta mai-le liang)	tiao	yu	
	(She bought two)	CLASS	fish	‘She bought two fish.’
	Word category:	classifier	noun	
	Number of choices:	small	large	
	Probability per word:	high	low	
	Information load:	low	high	

Each Chinese noun has a designated classifier when it occurs with a numeral. There are only a dozen or two common classifiers. In the classifier position, therefore, the number of choices is low, but in the noun position the number of choices is very high. Consequently, the probability of a classifier is high and its information load is low, whereas the probability of a noun is low and its information load is high.

Chao (1961) and Yuan (1999) have discussed the relation between the number of word choices and their information load. Ladd (1980a) has discussed the relation between information and stress, although he only focused on compounds. I shall return to Ladd’s discussion below.

Before we end this section, let us consider one more question: why should a word with more information be spoken with more stress? There are two possible answers. The first was suggested to me by LU Bingfu and YUAN Yulin independently, according to which the Information-Stress Principle reflects an intentional and functional behavior: the speaker puts more stress on words with more information in order to express ideas more clearly, so that the listener can understand the meaning better. This explanation in turn can be related to the Cooperative Principle of conversation of Grice (1989). However, while the explanation seems quite plausible, there are some questions. First, it is a hypothesis, and there is no evidence for it; indeed, it is hard to collect relevant evidence for it. Second, apart from contrastive stress and focus, Chinese speakers generally lack intuitive judgment on where phrasal stress is; the same can be said about English speakers, too. Therefore, it is hard to assume that speakers are consciously manipulating phrasal stress to improve communication. Third, when one is speaking to no audience (e.g. to one self), as it happens sometimes, there is no issue of the Cooperative Principle, yet the stress pattern still follows the Information-Stress Principle. Finally, the Cooperative Principle does not explain how information load is defined, and therefore, it offers no mechanism to assign stress properly. We still need to supplement it with Information Theory, and the application of Information Theory to word categories and syntactic structures.

The second explanation of the Information-Stress Principle is that it reflects a natural and subconscious behavior: when one thinks of words with more information, the relevant neurons are more excited, which in turn leads to greater stress in the articulation. This is similar to laughter: when one is happy, laughing is first of all an instinct, instead of the desire to let others know that you are happy. Similarly, the basic reason we close our eyes when we sleep is first of all an instinctive or natural behavior, instead of wanting to tell others that we are tired.

5. Using the Information-Stress Principle to explain stress effects

In this section I illustrate how the Information-Stress Principle can account for various kinds of stress effects.

5.1. The Non-head Stress of Duanmu (1990)

We begin with the Non-head Stress rule, because many other effects can be derived from it.

According to the Non-head Stress rule, any syntactic unit has a head and a non-head, represented as [X XP] or [XP X], where X is the syntactic head and XP the syntactic non-head. But why should XP have more stress? The answer lies in the Information-Stress Principle, detailed in (24).

(24) The Information-Stress Principle and Non-head Stress

	X (head)	XP (non-head)
Level	word	phrase
Choices	limited	unlimited
Probability	high	low
Information	low	high
Stress	low	high

The syntactic head is a unit at the word (or morpheme) level, which cannot be expanded. Regardless of how many choices there are, the number is limited. In contrast, the syntactic head is a unit at the phrase level, which can be expanded freely, and its number of choices is in principle unlimited. Therefore, the syntactic head has a high probability of occurrence and a low information load. In contrast, the syntactic non-head has a low probability of occurrence and a high information load. Therefore, the syntactic non-head has more stress.

5.2. The Compound Stress Rule of Chomsky and Halle (1968)

Chomsky and Halle (1968) propose that in a compound the first word has more stress. In the present analysis, this is because the first word is the syntactic non-head, which should have stress. As just discussed, the Non-head Stress is itself derivable from the Information-Stress Principle.

5.3. The Nuclear Stress Rule of Chomsky and Halle (1968)

Chomsky and Halle (1968) propose that in phrases the second constituent has more stress. In the present analysis, this is because in English the second constituent is the syntactic non-head. Consider the examples in (25), where underline indicates stress.

(25) *buy cars*
in school

According to Non-head Stress (which derives from the Information-Stress Principle), the syntactic non-head should be stressed. Therefore, stress goes to the right in such phrases.

In some cases, Non-head Stress offers a simpler analysis than the Nuclear Stress Rule of Chomsky and Halle (1968). Consider the examples in (26).

(26) Nuclear Stress: *Mary saw a rabbit*
Non-head Stress: *Mary saw a rabbit*

The Nuclear Stress Rule assigns stress to *rabbit* only. In contrast, Non-head Stress assigns stress to both *Mary* and *rabbit*, because they are both syntactic non-heads. Specifically, *a rabbit* is a determiner phrase, where the head is *a* and the non-head is *rabbit*. The verb *saw* is the head of the verb phrase. The sentence is either a tense phrase or an inflection phrase, where the head is either a tense or inflection element, and the subject *Mary* is a non-head.

In (26) *Mary* indeed has stress, as Non-head Stress predicts. In order to obtain the same result, the Nuclear Stress Rule must be supplemented with another rule, which is called the Stress Equalization Convention by Halle and Vergnaud (1987). Its function is to add stress to words that do not lie at the right end of a phrase. In this regard, the Non-head Stress analysis is simpler, because it obtains the result directly, without the Stress Equalization Convention.

5.4. The cumulative stress of Cinque (1993)

In the cumulative stress analysis (Cinque 1993), the deeper a syntactic unit is, the more stress it has. I have discussed earlier that cumulative stress is the same as non-head stress. The reason is that a syntactic non-head can expand, and therefore it can go deeper. In contrast, a syntactic head cannot expand and so it cannot go deeper. As a result, syntactic non-heads are generally deeper than syntactic heads.

5.5. The focus stress of Zubizarreta and Vergnaud (2000)

There are two parts in the stress theory of Zubizarreta and Vergnaud (2000). When the entire sentence is the focus, the stress assignment is similar to that of Non-head Stress. When the focus is smaller than the entire sentence, we first exclude the non-focus part, and then apply stress assignment to the remainder, which is again similar to the analysis of Non-head Stress. The question to explain then is: Why should we exclude the non-focus part from stress assignment?

By definition, the non-focus part is what has just been mentioned, or what is already known. According to the Information-Stress Principle, this part has no new information, or low information load. Therefore, it is not have given stress.

5.6. Stress differences among compounds

As discussed earlier, according to the Compound Stress Rule of Chomsky and Halle (1968), the first word of a compound should be stressed. However, there are some exceptions in English, as shown in (27), where underline indicates main stress.

- (27) Madison *Street*
Madison Avenue
Madison Road
Madison Drive

All expressions are street names. However, when the second word is *Street*, main stress is on the first word, but when the second word is *Avenue*, *Road*, or *Drive*, main stress is on the right. Such differences are hard to explain from a purely syntactic point of view.

Ladd (1980a) proposes that compound stress should not be derived from syntax, but from semantics, or information load. When a word does not have much information, it should not be stressed, but when a word has a lot of information, it should be stressed. Among English street names, *street* is the most common and the most expected, and therefore it has the least information load. In contrast, *avenue*, *road*, and *drive* are less common street names, and

therefore they carry more information and should be stressed. Ladd's conclusion is that English has just one stress rule, which is to assign stress to the right, including compounds. Compound stress is switched to the left only when the word on the right is a common one and has little information, such as *street* for street names.

Ladd's proposal is consistent with the Information-Stress Principle and can be adopted. However, we do not need to follow his conclusion that the default compound stress is on the right. Instead, we can keep the traditional idea that compound stress is on the left, unless the word on the right has high information load, in which case stress is switched to the right. In other words, English still has one stress rule, which is Non-head Stress.

Although both Ladd and the present analysis assume the Information-Stress Principle, and both assume one stress rule, Ladd's analysis is more complicated. First, his right-headed stress is an extra assumption, because it cannot be derived from the Information-Stress Principle. In contrast, Non-head Stress is directly derivable from the Information-Stress Principle. Second, most English compounds, including newly created ones, have left-headed stress, such as *oatmeal*, *shoe store*, *pancake*, etc. If the default stress rule is right-headed, it is hard to explain such compounds. Ladd is aware of such cases and proposes that short compounds can be treated as simple nouns, and they have initial stress because most nouns do. But Ladd's proposal cannot explain why other short nominal expressions, such as *red car*, *old man*, *new book*, etc., rarely have initial stress. In other words, why are they not treated as nouns? The answer, obviously, is that compounds are different from phrases, and it is not just a matter of length. In Ladd's analysis, therefore, compounds and phrases follow different rules. In the present analysis, they follow the same stress rule.

5.7. Stress differences among word categories

Different word categories often have different stress. For example, pronouns are often unstressed. In addition, grammatical words, such as prepositions, articles, and classifiers, are usually unstressed. On the other hand, content words, such as nouns and verbs, usually have stress. Moreover, nouns are more likely to have stress than verbs. Let us see how such effects can be explained by the Information-Stress Principle.

Pronouns pose a problem for several stress theories, including the Non-Head Stress rule of Duanmu (1990), the Nuclear Stress Rule of Chomsky and Halle (1968), and the cumulative stress rule of Cinque (1993). Pronouns often occur either as the subject or as the object, which is a syntactic non-head and should have stress. Yet pronouns are usually unstressed.

The Information-Stress Principle can explain the apparently exceptional property of pronouns. The referent of a pronoun is often already obvious to the participants of a conversation. Therefore, it has low information load and does not get stress.

The lack of stress in grammatical words (prepositions, classifiers, articles, etc.) is due to two reasons. First, in current syntactic theory, grammatical words are syntactic heads, and according to Non-head Stress, they do not receive stress. The second reason is that there are far fewer function words than content words. For example, English has about ten prepositions, but thousands of nouns. Therefore, in a preposition position, the probability of each word is high and the information load is low. In contrast, in a noun position the probability of each word is low and the information load is high. Therefore, prepositions usually have less stress than nouns.

Both nouns and verbs are content words, but they also have a stress difference, which has been noted before (Ladd 1980b: 90-92, Hayes 1995: 376), but without satisfactory explanation. In the present analysis, there are two reasons for their difference. First, verbs usually occur as the

head of a verb phrase, and according to Non-head Stress, they do not receive phrasal stress. In contrast, nouns often occur as the subject or the object, which is a syntactic non-head, which is assigned stress by the Non-head Stress rule. The second reason is that there are more nouns than verbs in English. In particular, in the CELEX English lexicon (Baayen et al 1993), there are 4,909 noun morphemes but just 1,628 verb morphemes. Thus, the probability of occurrence of an average verb is three times that of a noun, and its information load is one third of that of the latter. That could explain why verbs usually have less stress than nouns.

5.8. Emphasis and contrast

Any word in a sentence can be spoken with stress in order to express special emphasis. An example is shown in (28), where underline indicates emphasis.

- (28) *She bought two new books.*
She bought two new books.

Such emphasis may imply a contrastive meaning. For example, the second sentence could imply that *she bought a book, not him*.

According to Non-head Stress (Duanmu 1990) or cumulative stress (Cinque 1993), phrasal stress is determined by syntax. Since the sentences in (28) are identical, they ought to have the same stress pattern. To explain emphasis, additional assumptions must be made. However, if stress is determined by the Information-Stress Principle, then no additional assumption is necessary: the word under emphasis or contrast has more information of interest, and therefore it has more stress.

5.9. Variability in phrasal stress

In an influential article, Bolinger (1972) makes the point that phrasal stress is unpredictable, unless 'you are a mind reader'. Similarly, Hayes (1995) points out that, whereas word stress is usually stable, phrasal stress is often variable. Such variability is expected by the Information-Stress Principle. Word stress is usually unrelated to meaning. In contrast, phrasal stress is based on meaning or information load. In addition, the information load of a word is not fixed but dependent on not only syntax but also the context and the knowledge of the participants of a conversation. Naturally, it can be variable,

5.10. Frequency effect

We discussed earlier that frequent words are more likely to undergo stress reduction than infrequent words (Fidelholtz 1975). Some examples are repeated in (29), which show that frequent words undergo vowel reduction because of stress reduction, while infrequent words do not.

- (29) Frequent: *information* [ɚ] astronomy [ə]
 Infrequent: *importation* [oɚ] gastronomy [æ]

Previous stress theories, such as Chomsky and Halle (1968), Duanmu (1990), Cinque (1993), and Zubizarreta and Vergnaud (2000), offer no explanation for such effects. Under the Information-Stress Principle, an explanation is available. Frequent words have a higher probability of occurrence and a lower information load. In contrast, infrequent words have a lower probability of occurrence and a higher information load. Therefore, infrequent words are more likely to be stressed.

The frequency of a word can vary from speaker to speaker. For example, Fidelholtz (1975) points out that *trombone* is a frequent word for trombonists, who tend to reduce the first vowel to [ə]. For non-musicians, *trombone* is an infrequent word, and the first syllable has a full vowel with secondary stress. Similarly, *Toronto* is a frequent word for people who live in that city, and the word is pronounced as [trɒnto], dropping the first vowel. For other people *Toronto* is an infrequent word and the pronunciation is [təɹɒnto], where the first vowel is kept.

5.11. Summary

The Information-Stress Principle has broad applications. From it we can derive compound stress, Non-head Stress, and other versions of phrasal stress. In addition, the Information-Stress Principle can explain other stress effects which previous stress theories cannot account for.

6. Stress and language typology

Linguistic theories must address two questions. One is the common properties among languages. The other is special properties of individual languages. A popular approach is to adopt the principles-and-parameters approach of Chomsky (1981), according to which variation across languages can be accounted for in terms of a set of parameters. For concreteness, let us consider how parameters account for language variation in two areas: syllable structure and foot structure.

Blevins (1995: 219) proposes a set of parameters to account for variation in the size of the maximal syllable in different languages. The main parameters are shown in (30), along with their settings for English.

(30) Syllable parameters (Blevins 1995: 219)	
Parameter	English setting
Can the onset have two sounds?	yes
Can a diphthong fill the nucleus?	yes
In the coda allowed?	yes
Can the coda have two sounds?	yes
Can an extra C occur word initially?	(yes)
Can an extra C occur word finally?	yes

To account for possible and impossible onset clusters and coda clusters, additional parameters can be proposed, which we omit.

Parameters for foot structure have also been proposed (e.g. Prince 1983, Halle and Vergnaud 1987, and Hayes 1995). Some are listed in (31).

- (31) Foot parameters:
- a. Structure
 - i. Length: At most one syllable/two syllables/three syllables/unlimited
 - ii. Weight-sensitivity: a heavy syllable can/cannot be unstressed

- iii. Head: the head (stress) is on the left/right of a foot
- iv. Head property: a stressed syllable must/need not be heavy
- b. Direction of foot construction: from left to right/right to left
- c. Iteration: foot construction is/is not iterative
- d. Edges: skip the first sound/syllable/foot/word at the left/right of a foot/word/phrase

The parameter approach to language typology is wide spread. In Optimality Theory (Prince and Smolensky 1993), parameter settings are replaced by constraint rankings, but the perspective essentially remains the same: languages fall into different types, which can be predicted by different parameter settings or different constraint rankings.

If the Information-Stress Principle is correct, there is at least one area where there is no language typology, which is phrasal stress. In addition, Duanmu (2009) has argued that, as far as the maximal syllable size is concerned, there is no typology either. Therefore, linguistic theory may turn out to be simpler than previously conceived.

7. Can phrasal stress differ in different languages?

If the information-Stress Principle is universal, one would expect phrasal stress to work the same way in all languages. To many people this claim would seem too strong. In particular, there are languages are thought to have no stress. In addition, for languages that have stress, it is often thought that each can choose its own phrasal stress rules, different from those in others.

In some languages, such as English, it is possible to obtain a reasonable amount of agreement on phrasal stress. In other languages, such as Chinese, Japanese, Korean, and Thai, native judgment on stress is harder to obtain. Naturally, there is a lack of studies on stress in the second group of languages and many linguists thought they have no stress. For example, in a typological study of stress, Hyman (1977) surveyed 444 languages, and over 100 of them were thought to have no stress or accent, including Chinese. Similarly, Selkirk and Shen (1990) propose that Shanghai Chinese has no stress. However, I have argued elsewhere that both Standard Chinese and Shanghai Chinese have stress (Duanmu 1999a, 1999b), and that the lack of native judgment on stress in some languages is due to the fact that the pitch contour has been used for other purposes, in particular lexical tones or pitch accents. I shall also discuss a specific case in section 8, which is Japanese.

Let us now consider to what extent phrasal stress can vary across stress languages. Ladd (1996) argues that phrasal stress is not determined by universal principles; instead, each language can choose its own phrasal stress rules. As a result, in similar syntactic structures, the location of stress can differ from language to language. Let us consider some examples that Ladd offers.

First, consider Russian and English. In some Russian question, main stress falls on the verb, not the object. Two examples are shown in (32), where uppercase indicates main stress.

- (32) Russian questions: main stress on the verb
- a. Ona SPIT? 'she SLEEPS?'
 - b. Ona KUPILA knigu? 'she BOUGHT book?'

English questions: main stress on the verb or the object

- a. *Did she SLEEP?*
- b. *Did she buy a BOOK?*

The examples seem to show that Russian and English have different phrasal stress rules. However, in these examples the languages do not have the same syntax. For example, in English the auxiliary verb is fronted, whereas in the Russian sentence there is no fronting of an auxiliary verb. It can be shown that, if the syntax is different, main stress can be different, even if the meaning remains similar. Consider the examples in (33).

- (33) *You saw WHAT?*
What did you SEE?

The two sentences have similar meanings but differ in syntax. They also differ in main stress, in that it is on the object in one case and on the verb in the other case. Therefore, we should not compare sentences that have different syntax.

Ladd's second set of examples compares Rumanian and English. In Rumanian, the main stress of each clause falls on the same word. In contrast, in the corresponding English sentence, the main stress falls on different words in each clause.

- (34) Rumanian:
 [...o sa vedem] ce AVETI, si ce nu AVETI
 [...we'll see] what YOU.HAVE and what no YOU.HAVE

English:
 [...we'll see] *what you HAVE, and what you DON'T have.*

In the sentences, Rumanian and English indeed differ in what main stress falls. On the other hand, the two sentences do not have the exactly same syntax. Therefore, we do not expect them to have the same stress pattern.

Ladd's third set of examples concerns English itself. Sentences that seem to have similar syntax can differ in main stress: it falls on the subject in (35) but on the verb in (36).

- (35) Main stress on the subject: non-action
 a. *My UMBRELLA broke.*
 b. *The SUN came out.*
 c. *His MOTHER died.*
- (36) Main stress on the object: action
 a. *My brothers are WRESTLING.*
 b. *Jesus WEPT.*
 c. *The professor SWORE.*

However, as Ladd points out, the two groups of sentences differ in the property of the verb. The verbs in (35) often have the meaning of appearance or disappearance, and those in (36) are often actions. In addition, appearance verbs can appear before the subject, as seen in (37) and (38).

- (37) English appearance verbs
Out comes the sun.
There dies the tax bill.

- (38) Chinese appearance verbs
- a. huai-le yi-ba san
broke an umbrella
'An umbrella broke'
 - b. chu taiyang le
appear sun ASPECT
'the sun appeared'
 - c. hen zao si-le ma
very early dies mother
'(his) mother died very early'

In contrast to the verbs in (37) and (38), action verbs cannot occur before their subjects. Therefore, the two sets of verbs may have different syntax after all, and they do not constitute evidence for the claim that phrasal stress is independent of syntax.

8. Stress in Japanese

In Japanese every word has a specific pitch contour. In addition, to predict the pitch contour correctly, we only need to mark one specific location in the word, which is the last H-toned mora, after which the pitch would be L. This mark is called the accent (McCawley 1968). Phonetic studies also show that the accented syllable has greater prominence (Pierrehumbert and Beckman 1988, Kubozono 1993). It is reasonable to assume then that the accented syllable is a stressed syllable, although its phonetic realization is not entirely the same as that of a stressed English syllable.

However, as discussed in Duanmu (2008), Japanese poses at least two problems for the Information-Stress Principle. First, about 50% of the words in Japanese have no accent (Yokoyama 1979), in the sense that the pitch remains H through the end of the word, even when suffixes are added. The lack of stress in so many words poses a problem for the Information-Stress Principle, according to which we expect at least all content words, such as nouns and verbs, to have stress. Also, in a verb phrase, the syntactic head is the verb, and we expect it to have less stress than the object (the syntactic non-head). However, if the verb has an accent and the object noun does not, we get an unexpected stress pattern: the verb has stress but the object noun does not, contradicting the prediction of Non-head Stress.

A second problem is that in a compound noun, the first word (the syntactic non-head) is unaccented, whereas the second word (the syntactic head) is accented. This again contradicts the prediction of the Information-Stress Principle, according to which the first word should have more stress. An example is shown in (39), where the accented vowel is followed by an apostrophe.

- (39) Compound stress in Japanese (traditional analysis)
- | | |
|-------------------|---------------|
| [+A]-[-A] | [-A]-[+A] |
| a'Npo-jooyaku → | aNpo-jo'oyaku |
| 'security treaty' | |

[+A]-[+A] [-A]-[+A]
 sha'kai-shu'gi → shakai-shu'gi
 'society principle (socialist)'

[-A]-[-A] [-A]-[+A]
 kyooiku-seesaku → kyooiku-se'esaku
 'education-policy'

In the first compound, 'security' has an accent ([+A]), which is lost in the output ([-A]), and 'treaty' has no accent but gains one in the output. In the second compound, both words have an original accent, but the first word loses it in the output. In the third compound, neither word has an original accent, but the second word gains one in the output. In other words, regardless of the original accents, a compound word must have an accent on the second part and no accent on the first part, in contradiction to the prediction of Non-head Stress.

A solution to the problems is proposed by Duanmu (2008), according to which every Japanese word has at least one accent, and some have two. The first accent, if present, is invariably on the first syllable. The location of the second accent, which is the same as the traditional accent, is variable and must be lexically specified. The compounds in (39) are re-analyzed in (40), where accent A is indicated by a double apostrophe and accent B by a single apostrophe.

(40) Compound stress in Japanese (Duanmu 2008)

B-A A-B
 a'Npo-jo''oyaku → a''Npo-jo'oyaku
 'security-treaty'

B-B A-B
 sha'kai-shu'gi → sha''kai-shu'gi
 'society-principle (socialist)'

A-A A-B
 Kyo''oiku-se''esaku → kyo''oiku-se'esaku
 'education-policy'

In the new analysis, every word has an accent. In addition, it is possible to claim that compound stress is assigned to the left (see Duanmu 2008). Finally, there is an explanation why the compound stress is A-B: if a compound is treated as a word, it should have a word stress pattern, and the largest word stress pattern is A-B.

9. Conclusions

Although word stress can differ from language to language, stress above the word level (phrasal stress) seems to be similar across languages. Every language has phrasal stress, and every language follows the same principle for phrasal stress, which is the Information-Stress Principle. According to the principle, words with high information load should be stressed and words with low information load need not. The principle can explain a range of stress effects, including

regular phrasal stress, contrastive stress, the relation between stress and syntax, and the relation between stress and word frequency.

If the present study is correct, there is no typological variation, or the need for parameters, in at least some areas of language. The result simplifies language typology and linguistic theory.

* A draft of this paper was presented at the Third Symposium of Overseas Chinese Linguists in June, 2007. I thank the conference hosts, Professor Yang Yeming, the Linguistics Institute of Xuzhou Normal University, and the Yuyan Kexue editorial office, and the organizers at Xuzhou Normal University and Xinjiang University, for their invitation and hospitality. I also thank the audience and LU Bingfu for their comments. This paper was originally written in Chinese and appeared in the journal Yuyan Kexue in 2007. In the translation I have added a section on stress in Japanese.

References:

(Note: Chinese reference names are given in brackets after the English name)

- Baayen, R. Harald, Richard Piepenbrock, and L. Gulikers. 1993. The CELEX lexical database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Blevins, Juliette. 1995. The syllable in phonological theory. In *The handbook of phonological theory*, ed. John Goldsmith, 206-244. Cambridge, MA: Blackwell.
- Bolinger, Dwight. 1972. Accent is predictable (if you are a mind reader). *Language* 48.4: 633–644.
- Chao, Yuen-Ren. 1961. Degree of meaningfulness in language constituents. *Tsing Hua Journal of Chinese Studies* 2.2: 1-17.
[赵元任, 1961, 語言成份裏意義有無的程度問題, 《清华学报》 2.2: 1-17.]
- Chomsky, Noam. 1971. Deep structure, surface structure and semantic interpretation. In *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology*, ed. Danny D. Steinberg & Leon A. Jakobovits, 183-216. Cambridge, UK: Cambridge University Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. N.Y.: Harper and Row.
- Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24.2:239-297.
- Duanmu, San. 1990. *A formal study of syllable, tone, stress and domain in Chinese languages*. Doctoral dissertation, MIT, Cambridge, Mass.
- Duanmu, San. 1999a. Metrical structure and tone: evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8.1: 1-38.
- Duanmu, San. 1999b. *Zhongyin lilun he Hanyu de ci chang xuanze* [Metrical theory and word length choices in Chinese]. *Zhongguo Yuwen* 1999.4 (271):246-254.
[端木三, 1999b, 重音理论和汉语的词长选择, 《中国语文》1994, 4 (271) : 246-254.]
- Duanmu, San. 2005. The tone-syntax interface in Chinese: some recent controversies. In *Proceedings of the Symposium “Cross-Linguistic Studies of Tonal Phenomena, Historical Development, Tone-Syntax Interface, and Descriptive Studies”*, December 14-16, 2004, ed. Shigeki Kaji, 221-254. Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.

- Duanmu, San. 2008. A Two-Accent Model of Japanese word prosody. *Toronto Working Papers in Linguistics* 28: 29-48.
- Duanmu, San. 2009. *Syllable Structure: the limits of variation*. Oxford: Oxford University Press.
- Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. *CLS* 11, 200-213.
- Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27:765-768.
- Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and Speech* 1:126-152.
- Grice, Paul. 1989. *Studies in the Way of Words*. Cambridge MA: Harvard University Press.
- Halle, Morris, and Jean-Roger Vergnaud. 1987. Stress and the cycle. *Linguistic Inquiry* 18:45-84.
- Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.
- Hyman, Larry. 1977. On the nature of linguistic stress. In *Studies in stress and accent*, ed. Larry Hyman, 37-82. *Southern California Occasional Papers in Linguistics* 4. Los Angeles: Department of Linguistics, University of Southern California.
- Kubozono, Haruo. 1993. *The organization of Japanese prosody*. Tokyo: Kurocio Publishing.
- Ladd, Dwight Robert. 1980a. English Compound Stress. *Proceedings of the Eleventh Meeting of the Northeast Linguistic Society (NELS 11)*, 173-189. GLSA Publications, Department of Linguistics, University of Massachusetts, Amherst.
- Ladd, Dwight Robert. 1980b. *The Structure of intonational meaning: evidence from English*. Bloomington: Indiana University Press.
- Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge University Press.
- Liu, Fang, and Yi Xu. 2005. Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetics* 62: 70-78.
- McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.
- Pierrehumbert, Janet, and Mary Beckman. 1988. *Japanese tone structure*. Cambridge, Mass.: MIT Press.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365-424.
- Prince, Alan. 1983. Relating to the grid. *Linguistic Inquiry* 14.1: 19-100.
- Prince, Alan, and Paul Smolensky. 1993. *Optimality theory: Constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado.
- Ritter, Elizabeth. 1991. Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. In *Perspectives on Phrase Structure: Heads and Licensing*, ed. Susan D. Rothstein, 37-62. *Syntax and Semantics* 25. San Diego: Academic.
- Selkirk, Elisabeth, and Tong Shen. 1990. Prosodic domains in Shanghai Chinese. In *The phonology-syntax connection*, ed. Sharon Inkelas and Draga Zec, 313-337. CSLI, Stanford University, Stanford, Calif. Distributed by University of Chicago Press.
- Shannon, Claude E. 1948. The mathematical theory of communication. *Bell System Technical Journal* 27: 379-423 and 623-656, July and October.
- Shen, Jiong. 1985. Beijinghua shengdiao de yinyu he yudiao [Pitch range of tone and intonation in Beijing dialect]. *Beijing Yuyin Shiyanlu* [Working Papers in Experimental Phonetics], ed. by Tao Lin and Lijia Wang, 73-130. Beijing: Peking University Press.
- [沈炯 1985, ‘北京话声调的音域和语调’, 《北京语音实验录》, 林焘, 王理嘉等著, 73-130. 北京: 北京大学出版社.]

- Wang, Zhijie, and Shengli Feng. 2006. Tonal contrast and disyllabic stress patterns in Beijing Mandarin. *Yuyan Kexue* [Linguistic Science] 5.1: 3-22.
[王志洁, 冯胜利, 2006, 声调对比与北京话双音组的重音类型, 《语言科学》, 第五卷 第一期 (总第 20 期), 3-22.]
- Yokoyama, Shoichi. 1979. "zisyo midasigo no akusento no seikaku ni tuite" [On the nature of the accent of dictionary-listed words]. *Nihon Onkyoogakkai Kooen Ronbunshuu*. The Acoustical Society of Japan.
- Yuan, Yulin. 1999. A cognitive explanation of the order of modifiers and its implication. *Social Science in China* 1999.2: 185-201.
[袁毓林, 1999, 定语顺序的认知解释及其理论蕴涵, 《中国社会科学》第 2 期, 185-201 页.]
- Zubizarreta, Maria Luisa, and Jean-Roger Vergnaud. 2000. Phrasal Stress and Syntax. In *Progress in Grammar* (an electronic book <http://www.roquade.nl/meertens/progressingrammar/index.html>), ed. Marc van Oostendorp and Elena Anagnostopoulou. Publisher: Roquade, Amsterdam /Utrecht /Delft.