New evidence against the modularity of grammar: Constructions, collocations, and speech perception

MARTIN HILPERT*

Abstract

This paper combines quantitative corpus data and experimental evidence to address the question whether speech perception is influenced by knowledge of grammatical constructions and, more specifically, knowledge of preferred collocation patterns of these constructions. Lexical identification tasks are devised in which subjects are presented with synthesized, phonetically ambiguous stimuli. The results suggest that knowledge of constructions and collocations influences speech perception, thus providing evidence for a usage-based, non-modular view of grammar.

Keywords: *modularity of grammar; constructions; collocations; lexical identification task; phonemic boundaries; compensation for coarticulation.*

1. Introduction

Usage-based approaches to language (Barlow and Kemmer 2000, Bybee and Hopper 2001, Bybee 2006) hold that repeated usage events over time shape grammar. One foundational aspect of this hypothesis is the often-made observation that frequent words tend to reduce phonetically and phonologically (Zipf 1935; Hooper 1976; Bybee 2000, 2001, *inter alia*). Several phenomena empirically support this point. Jurafsky et al. (2001) find that word-final t/d deletion correlates positively with the relative

* I would like to thank Katherine Crosswhite, Suzanne Kemmer, Nancy Niedzielski, Joan Bybee, Anatol Stefanowitsch, and two anonymous referees for *Cognitive Linguistics*, who have offered helpful comments on earlier versions of this paper. Also, the audience of the 8th CSDL in San Diego provided valuable suggestions. The usual disclaimers apply.
Please address correspondence to ⟨hilpert@icsi.berkeley.edu⟩.

Cognitive Linguistics 19–3 (2008), 483–503 0936–5907/08/0019–0483
DOI 10.1515/COGL.2008.018

frequency of lexical items such as *want* or *mind*. Similarly, Cooper and Paccia-Cooper (1980) observe increased palatalization of word-final stops before a glide in items with higher frequency. Palatalization is thus more likely in a phrase such as *did you*, as compared to *mind you*. Finally, Gregory et al. (1999) report that word duration is relatively shorter for items with higher text frequency and greater contextual probability. These and many other studies strongly support the relation of frequency and phonetic reduction, and hence the usage-based model.

The present study focuses on another tenet that is commonly held in usage-based approaches and elsewhere, but which as yet has not been sufficiently supported through empirical studies. A core assumption in both cognitive linguistics (Langacker 1987) and connectionist modeling (McClelland et al. 1986) has been that the mental representation of grammar is non-modular. The common distinction between a syntactic module, the mental lexicon, and a phonological module is rejected in favor of a monotonic structure of grammar. In many formalist frameworks, a modular organization of grammar with particular emphasis on the autonomy of syntax is presupposed, following suggestions and definitions of Fodor (1983). To illustrate, Newmeyer (1998: 23) defines the autonomy of syntax in terms of the following hypothesis:

> The Autonomy Of Syntax (Autosyn):
> Human cognition embodies a system whose primitive terms are nonsemantic and nondiscourse-derived syntactic elements and whose principles of combination make no reference to system-external factors.

This hypothesis does of course not deny that information from different grammatical modules is integrated at some level of linguistic processing, or even at multiple levels. The crux of the argument is therefore that certain types of information are processed in one module but disregarded in another. Frazier (1987) lays out how, for instance, acoustic spectra are instrumental for parsing speech into words, whereas the resolution of anaphoric reference seems quite unrelated to this particular task. Clifton (1991: 97) explicates this assumption in the following quote:

> Modules are defined, in part, in terms of the information relevant to them, and thus in terms of their representational vocabularies. Information about letters or speech sounds is relevant to the lexical module (if there is such a thing), but is of no possible value to the syntactic module.

In cognitive linguistics, the idea of modularity has been repeatedly criticized (Bybee 2006; Fillmore et al. 1988; Goldberg 1995, 2006; Lan-

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 484)

gacker 1987). The alternative view is that lexical and syntactic knowledge are stored in the same way, and are not processed by different mental modules. A representative position is expressed by Langacker (2005):

> Lexicon, morphology, and syntax form a continuum, divided only arbitrarily into discrete "components". Everything along this continuum is fully describable as assemblies of symbolic structures. A symbolic structure is specifically defined as the pairing between a semantic structure and a phonological structure (its semantic and phonological poles).

While this approach denies the existence of distinct modules that handle different aspects of grammatical knowledge, little experimental evidence has been offered to demonstrate the uniformity of grammar. As a notable exception, Tanenhaus et al. (1995) demonstrate that visual information has an immediate influence on syntactic processing: A sentence such as *Put the apple on the towel in the box* presents hearers with a temporary syntactic ambiguity, as the prepositional phrase *on the towel* can initially be understood as a destination. If hearers are given a visual context that contains an apple, a towel, and a box, they will, for a brief period, pay close attention to the towel. Tanenhaus and colleagues show that the situation is very different if the visual context contains a second apple that is placed on a napkin. When presented with a contrastive set of two apples, a towel, and a box, hearers do not initially parse *on the towel* as a directional prepositional phrase—little attention is paid to the irrelevant sole towel. Tanenhaus and colleagues interpret this result as evidence against the modularity of syntactic processing.

Despite these findings, the hypothesis that grammar is non-modular is still both more speculative and not as well supported as the more general hypothesis that grammar is shaped through usage. The present study addresses the need to demonstrate more thoroughly that knowledge of language is indeed a large inventory of symbolic pairings of sound and meaning, and nothing else.

The issue of modularity is closely related to the question of how auditory perceptual input is integrated with knowledge of lexical items and syntactic structures. Is speech perception a purely bottom-up process in which sounds are sequentially parsed into phonemes, words, phrases, and sentences, or are there lexical and syntactic top-down effects that guide the perception of speech? Strict bottom-up organization would accord with (though not necessitate) a modular approach to grammar. Conversely, top-down effects on speech perception, in which lexical or syntactic levels of processing interact with the processing of speech sounds, are more naturally accounted for in a non-modular approach. It has to be

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 485)

pointed out, though, that both types of organization could in principle be modeled by either a modular or non-modular architecture.

Lexical top-down effects on speech perception have been reported on several occasions (Elman and McClelland 1988; Ganong 1980; Magnuson et al. 2003; Warren and Warren 1970). As will be explained in more detail below, these effects only operate on lexical units and therefore do not bear on the question of the autonomy of syntax. The present study goes beyond lexical effects and presents a top-down effect on speech perception that is driven by speakers' knowledge about constructions and non-lexicalized collocations. The fact that this type of knowledge has an effect on the perception of auditory input demonstrates the immediate interrelatedness of syntax and phonology, and thus constitutes new evidence against the purported modularity of grammar.

Recently, empirical evidence for syntactic effects on speech production has been presented by Gahl and Garnsey (2004). In a study of pronunciation variation, they show that words are not only shortened if their overall text frequency is high, but also when their syntactic context makes them highly probable. The overall duration of same word is shorter in contexts where it is more likely to occur, and hence more easily identified by the hearer.

To illustrate, in the corpus data used by Gahl and Garnsey, the verb *suggest* co-occurs with sentential complements more often than with direct objects, such that sentences like *The director suggested the scene should be filmed at night* are more likely than *The director suggested the scene between Kim and Mike* (2004: 752). In a production study that measures reading times, Gahl and Garnsey find that syntactic biases toward one complementation pattern significantly correlate with reduced production of the verb in question (2004: 763). Verbs such as *argue*, *believe*, *claim*, *conclude*, *confess*, or *decide* are pronounced shorter if they occur with a sentential complement, and longer if they occur with a direct object. Conversely, verbs such as *accept*, *advocate*, *confirm*, or *emphasize* are reduced when occurring with a direct object, but not when they take a sentential complement. Gahl and Garnsey conclude that the probabilities of different complementation patterns are mentally represented for each verb, and that this knowledge of syntactic probabilities affects speech production (2004: 768).

These findings suggest that the organization of grammar is in fact non-modular. On a strictly formalist view, the relative frequencies of syntactic patterns would not be part of any grammatical module to begin with, as it is held that usage does not affect the mental representation of grammar (Newmeyer 2003). On any modular view, the fact that syntactic representations affect subphonemic speech production would require an elaborate interface between modules, effectively reducing the autonomy of each respective module to a relative degree. The only conceivable explanation in terms of strict modularity would require the ad-hoc postulation of different lexical entries for verbs such as *suggest*, only differing in relative length, depending on their complementation patterns. Since such a solution requires a fair amount of technical machinery and auxiliary assumptions, Gahl and Garnsey point out that "the most parsimonious accounts of these effects will be ones in which the grammar itself is enriched with probabilistic information" (2004: 769).

The present study aims to provide further evidence for the non-modular view of grammar, and extends the approach of Gahl and Garnsey to another domain of language use: it will be argued that syntactic probabilities affect not only speech production, but also speech perception.

2. The experimental paradigm

The experimental paradigm used in the present study is that of lexical identification tasks. Subjects hear a stimulus and are asked to identify the word or phrase they perceived by selecting an orthographical representation on a computer screen. The stimuli used in this task are often not unambiguously identifiable. The stimuli are rendered ambiguous through synthesized elements that lie on a continuum between two phonemic poles, such as /p/ and /b/. For minimal pairs such as *pear* and *bear*, the intermediate steps on the continuum allow two possible interpretations, each of which is a lexical word of English. The experiment determines at which step of the continuum subjects flip from one interpretation to the other.

One of the best known applications of this experimental paradigm is the demonstration of the so-called categorical perception of speech sounds by Liberman et al. (1957). Liberman and colleagues showed that speech perception differs fundamentally from other types of perception. For instance, if subjects are presented with a color continuum from red to orange, that continuum is perceived as a gradual change that goes through a stage of orange-red in the middle. By contrast, if subjects are presented with a continuum from the syllable /pa/ to the syllable /ba/, each stimulus is perceived as either one or the other. No stimulus is perceived as midways between /pa/ and /ba/. This finding suggests that speech perception is categorical and depends on a specialized speech processing system, but this point has been subject to controversy (Fry et al. 1962, Kewley-Port and Luce 1984). This particular debate is not of

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 487)

concern here, as the present study merely shares the experimental paradigm, not the theoretical stakes, with these early studies.

An aspect of speech perception that is relevant to the present study has been discussed by Warren and Warren (1970), who observe a remarkable effect: When sounds are cut out from a recording and replaced by a non-phonemic sound such as a cough, hearers will "fill in" the missing sound without even being aware of it. For instance, if the first /s/ is replaced in a recording of the word *legislatures*, hearers will fail to hear that the word has been altered, even when they are explicitly asked to identify the spliced element. Warren and Warren call this effect *phonemic restoration*.

They further find that phonemic restoration is sensitive to the meaning of the context. In a sentence such as *It was found that the ⟨cough⟩ eel was on the shoe*, hearers robustly restore the word *heel*. By contrast, replacing the last word of the sentence with *axle* leads to the restoration of *wheel*. The restored element appears thus to be the semantically most appropriate candidate from a set of phonologically related items.

Another application that is similar in spirit to the present investigation was developed by Ganong (1980). Ganong found that hearers perceive phonologically ambiguous stimuli with a lexical bias: if a string of phonemes can be interpreted as a lexical word, subjects will favor this interpretation over a competing interpretation that is merely a phonotactically legal non-word. To illustrate, a stimulus that is ambiguous between *face* and *faish* will tend to be perceived as *face*. The Ganong-effect can be measured by comparing responses to different continua of stimuli. Responses to a continuum from *face* to *faish* will differ markedly from responses to a continuum from /ɛs/ and /ɛʃ/, which apart from the missing onset is identical to the *face* to *faish* continuum, but which offers only non-word syllables as possible interpretations. In the first continuum, subjects will be biased towards the competitor interpretation that is a lexical item (*face*). In the second continuum, none of the competitors is a lexical item, and so comparatively fewer responses will identify an ambiguous stimulus as /ɛs/. This effect can be interpreted as a shift in the perceptual category boundaries of phonemes such as /s/ and /ʃ/.

While both phonemic restoration and the Ganong-effect represent striking lexical top-down effects in the processing of auditory input, they do not address the hypothesis that syntax is an autonomous module of grammar. In order to put this hypothesis to the test, the present study investigates whether not only lexical words, but also larger syntactic units such as constructions (Goldberg 2006) or collocations can affect speech perception. Under the notion of syntactic effects, the present approach subsumes even lexico-grammatical dependencies such as the collocational preferences of particular grammatical constructions (Stefanowitsch and

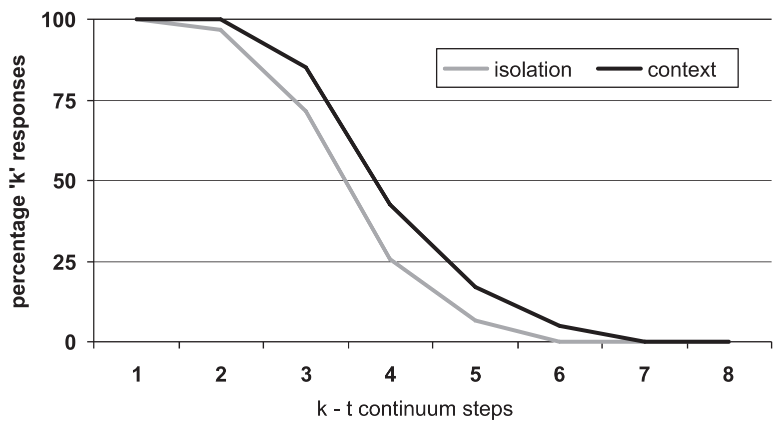

Figure 1. *Idealized curves of phonemic categorization*

Gries 2003). The present study devises lexical identification tasks in which subjects hear ambiguous stimuli that are embedded in different syntactic and collocational contexts. Unlike Ganong (1980), the present study uses stimuli for which both competing interpretations are actual words of English, as for example *cry* and *try*. Further, unlike Warren and Warren (1970), the present study employs stimuli whose competing interpretations are all semantically viable, i.e., hearers will not be given the biased choice of, say, a wheel being either on an axle or on a shoe. If a syntactic context significantly leads subjects to favor one competitor over the other, this constitutes evidence against the autonomy of syntax. The basic idea of the experiments in this study is thus to test whether syntactic context can shift category boundaries in phonemic perception. Figure 1 shows two idealized curves of phoneme perception in a continuum from /k/ to /t/.

The grey curve represents responses to stimuli heard in isolation, the black curve represents responses to stimuli heard embedded in syntactic context. The null hypothesis of the present study is that syntactic context has no influence whatsoever on phonemic speech perception. The stimuli should be categorized the same, whether they are heard in isolation or embedded within a given syntactic context. The two curves should thus coincide. The research hypothesis is that syntactic context can introduce a bias that leads subjects to shift their phonemic category boundaries, such that the two curves coincide at the ends of the continuum, but diverge in the middle. These alternative hypotheses are weighed in three different experiments.

The first experiment of this paper tests the general question whether syntactic knowledge has any effect on lexical identification. It is shown that constructions actually influence the perception of phonemically ambiguous stimuli. Given a stimulus that is ambiguous between two lexical elements in a construction, subjects are more likely to identify the stimulus as an element that frequently occurs in the respective construction, and less likely to identify it as an element that only sparsely occurs in that construction.

The second experiment tests how robust the findings of the first experiment are. If constructions influence the perception of phonemically ambiguous stimuli, it should be possible to find constructions that induce opposing biases, thus shifting the perceptual category boundaries in opposite directions. The results show that this is the case, corroborating the findings of the first experiment.

In a third experiment it is tested whether constructional context does not only influence the level of phonemic processing, but also extends to low-level phonetic processing. While the results of the first two experiments could be dismissed as operations that potentially involve the late re-categorization of 'misheard' words, the third experiment investigates whether syntactic knowledge directly and instantaneously influences lower levels of speech perception. The used test case is whether constructional context can trigger the phonetic effect of compensation for coarticulation (Elman and McClelland 1988). In processing naturally occurring speech, hearers accommodate the fact that any string of phonemes is affected by coarticulation. The production of every speech sound is influenced by its preceding elements that are coarticulated with it. For example, the /k/ in *this car* will sound somewhat different than the /k/ in *one car*. Compensation for coarticulation can be thought of as an increased "tolerance", such that even less than perfect examples of a phoneme are categorized as such, when hearers know that there is a reason for the undershoot. Compensation for coarticulation is necessarily a low-level phonetic process that works instantaneously to keep up with the natural speech flow. The results of the third experiment show that constructions can indeed induce compensation for coarticulation. This means that knowledge of syntax has an effect on phonetic processing, which in turn constitutes evidence against syntactic modularity of the kind hypothesized in Newmeyer (1998).

3. Materials and participants

The speech stimuli used in the present study are based on recordings of an adult female human voice, speaking in a standard American English

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 490)

variety. After the recordings, the stimuli were altered with a computerized synthesizer to yield ten-step continua between two phonemic poles, such as /k/ and /t/. The chosen method of synthesizing was sample-averaging. This technique divides two wave forms of the same length into small slices at the rate of 44.1 kHz and creates continua of ambiguous sounds by laying wave forms from the two different sources on top of each other. Depending on how strong each source is represented at a given continuum step, the resulting wave form sounds more or less like one of the original sources. The synthesized stimuli are embedded in unaltered recordings of actual words to yield a continuum between, say, the English words *cry* and *try*. The end points of the continuum are unambiguously perceived as *cry* and *try*, but the point at which the perceptual crossover from *cry* and *try* occurs will vary from person to person. In some experiments reported in this paper, the outer continuum steps were discarded if pilot studies indicated that even steps that lay more towards the center of the continuum were identified unambiguously across subjects.

Fifteen volunteer subjects with self-reported normal hearing, normal or corrected to normal vision, and English as their native language participated, each one in all three of the experiments. Since this was a procedure of about 45 minutes, subjects were instructed to take breaks whenever they felt the need for doing so. The experimental design was fully self-paced through mouse clicks and allowed for subject-controlled breaks. All subjects were Rice University undergraduate or graduate students that were either paid or given course credit for their participation. None of the data had to be excluded.

4. Experiment 1—The English *make*-causative

In the first experiment, subjects are presented with ambiguous speech signals within a construction that is intended to bias the lexical identification process towards one of the two competing interpretations. The construction used in this experiment is the English *make*-causative, which has a strong bias towards verbs of emotion and psycho-physiological reaction (Kemmer 2001). Typical examples are *It made me feel dizzy* or *That makes it look a lot bigger*; examples involving activity verbs such as *He made me do it* are much less frequent, despite the high text frequency of the verb *do*. Table 1 shows the twenty most frequent verbs from an exhaustive extraction of the make-causative construction from the British National Corpus (Leech 1992), which yields 10,708 examples.

While the verb *cry* occurs 73 times in the make-causative construction, the verb try, which is not shown in Table 1, occurs only eleven times. As a

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 491)

Table 1. *The 20 most frequent verbs in the English make-causative construction in the BNC*

Verb	Tokens	Verb	Tokens
feel	1654	appear	142
look	822	happen	119
think	542	come	111
laugh	358	realise	111
seem	293	see	100
work	264	pay	97
sound	258	meet	93
go	237	stand	91
want	195	take	74
wonder	157	cry	73

minimal pair with *cry*, it affords a test case for the effect of constructional context on speech perception. Note that *try* is ten times as frequent in discourse as *cry* (Francis and Kucera 1982), such that the frequency of *cry* and *try* in the make-causative is asymmetrical to their overall frequency. At any rate, the context of the make-causative should bias the categorization of stimuli ambiguous between *cry* and *try* towards *cry*. The carrier phrase that is used in the experiment is the phrase *They made me*, which is followed by a signal that ranges on a eight-step continuum from /trai/ to /krai/. It is hypthesized that the constructional carrier phrase biases hearers towards perceiving a principally ambiguous signal as /krai/. To test this hypothesis, subjects categorized the ambiguous signals both within the constructional frame and in isolation.

4.1. *Method*

4.1.1. *Materials.* For the experiment, a ten-step /t/-/k/ continuum was created using a sample-averaging script within Praat (Boersma and Weenink 2005). The source signals were the items *cry* and *try*, recorded from the speech of a female native speaker of American English. Input sections were selected that contained the burst of the consonant as well as the first four glottal pulses. Again using Praat, the longer one of the two sections was shortened such that both sections were of equal length. From these continuum endpoints, intermediate signals were created in 10% steps. Each of the resulting /t/-/k/ continuum steps was concatenated with the remaining stretch of the recording of *try*, which comprised the entire word minus the burst and the first four glottal pulses. This procedure yields a continuum of sounds, the first one an unambiguous /krai/, and the last one an unambiguous /trai/. In pilot studies the two

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 492)

endpoint steps were identified unambiguously and hence discarded; only the intermediate eight steps were used. For the carrier phrase, the phrase *they made me* was recorded from the same speaker, and subsequently concatenated with each of the eight stimuli. The subsequent concatenation of stimuli types yields 16 different stimuli (carrier phrase and null context by eight try-cry continuum steps).

4.1.2. *Procedure.* The experiment was conducted using PsyScope 1.2.5 (Cohen et al. 1993). Subjects were given on-screen instructions, stating that they would see a clickable red dot as a fixation point at the center of the computer screen. After clicking the dot, they would hear a prerecorded sound file and have to identify the percept as a word of English in a two-way choice. Orthographical representations in a 44 pt font were displayed to the left and right side of the screen. The same orthographical representation would appear in the same place throughout. For each subject, this experiment involved 64 trials, such that each of the eight continuum steps was heard eight times, four times in isolation, and four times in the constructional context. No filler trials were used.[1] Stimuli were presented in randomized order, while the relative positions of the orthographical representations were kept constant.

4.2. *Results*

Figure 2 summarizes the outcome of the first experiment. In both conditions, the perceptual crossover from /k/ to /t/ occurs between steps three and six. The outer two steps on either side of the continuum are unambiguously identified by all subjects. The figure shows that the categorization curve is drawn half a step towards the right side of the continuum in the context of the make-causative construction, which is consistent with the research hypothesis. More instances of ambiguous sounds are identified as *cry* if they are presented in the constructional carrier phrase. A repeated measures ANOVA was conducted for the *cry* responses in isolation and in the constructional context to measure the effect of the constructional carrier phrase. The calculation is based on all *cry* responses of fifteen subjects in two different conditions (isolation, causative) across the eight steps in the synthesized continuum.

1. To the extent that filler trials serve to obscure the research question in an experimental design, they were not deemed necessary in the present study. Additionally, since participants completed all three experiments in one lengthy sitting, fillers would have meant an additional strain on the participants.

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 493)

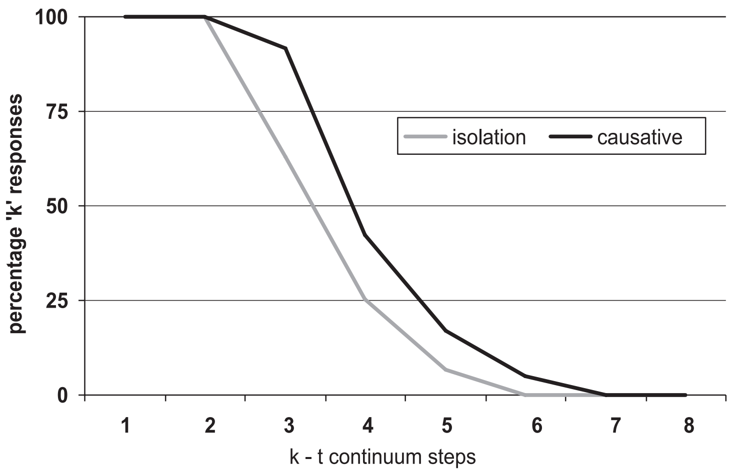

Figure 2. *Perception of the /trai/-/krai/ continuum in isolation and the causative*

The constructional effect is significant in a by-subject analysis ($F_{(1,14)} = 18.44$, $p < 0.001$) and approaches significance in the corresponding by-item analysis ($F_{(1,7)} = 4.17$, $p = 0.08$).

5. Experiment 2—Using collocations to induce opposing biases

While the first experiment is designed to merely explore whether constructional knowledge, which falls into the domain of syntax, has an effect on speech perception, the second experiment tests how robust this effect is. Given that only the by-subject analysis returned a significant result in the first experiment, further investigation seems necessary. Again, subjects are presented with phonologically ambiguous stimuli that are embedded in constructional frames. This time however, subjects hear the same stimuli in three different conditions. In the control condition, subjects hear the stimuli in isolation. In the two other conditions, the stimuli are embedded in collocations that exhibit different lexical preferences. The research hypothesis is that each carrier phrase biases lexical identification towards its lexical preference, while the control condition yields responses that fall in between the two other conditions. The null hypothesis is that all three conditions should receive similar responses, or responses that differ randomly.

The phrases *it's always* and *it's getting* are used as carrier phrases because they collocate heavily with the elements worse and worth, which form a phonological minimal pair of English. The collocational prefer-

Table 2. *The 20 most frequent items after* it's always *and* it's getting *in the BNC*

it's always	Tokens	it's getting	Tokens
the	74	a	25
a	64	late	20
like	14	dark	15
nice	14	more	11
difficult	13	worse	10
there	12	to	9
worth	11	on	8
easier	9	the	7
going	9	too	7
better	8	very	7

ences are opposed to each other: while *it's always* frequently occurs with *worth*, *it's getting* is frequently followed by *worse*. The reverse combinations are not ungrammatical, but very infrequent. Sentences such *as It's getting worth investing again* are thus rarely seen. Table 2 shows the ten most frequent items that occur after *it's always* and *it's getting* in the BNC. The table is based on 502 occurrences of *it's always* and 292 tokens of *it's getting*.

As expected, function words such as the determiners *the* and *a*, and prepositions such as *to* and *on* are among the most frequent elements. However, both lists also contain open-class elements such as the adjectives *nice*, *difficult*, *easier*, and *better* with *it's always*, and *late* and *dark* with *it's getting*. What matters to the present analysis is that the minimal pair members *worse* and *worth* approximate a complementary distribution across the two collocational environments. In terms of absolute frequency, *worse* and *worth* occur at the same order of magnitude (Francis and Kucera 1982), such that any observed effect should not be due to stronger familiarity with one of the two competitors.

The distributional asymmetry between the items after *it's getting* and *it's always* should lead subjects to interpret the same ambiguous stimuli in different ways, depending on the preceding context. The question pursued in this experiment is whether this difference is strong enough to induce opposing biases that are statistically significant, and that are both distinct from intermediate responses to the control condition.

5.1. *Method*

5.1.1. *Materials.* A ten-step /s/-/θ/ continuum was created using the same sample-averaging script within Praat. The source signals were the items *worse* and *worth*, which were recorded as spoken by a female native speaker of English. Input sections were selected that contained the last four glottal pulses from the vowel /□/. The longer /s/-section was shortened such that it was of equal length as the /θ/-section. From these continuum endpoints, intermediate signals were created in 10% steps. Each of the resulting /s/-/θ/ continuum steps was concatenated with the remainder of the word *worth*, which comprised the entire word minus the last four glottal pulses and the frication. This procedure yielded a continuum of sounds, the first one an unambiguous /w□s/, and the last one an unambiguous /w□θ/. The first points and the last two points of the continuum were interpreted unambiguously in pilot tests, such that they were discarded and not used in the actual experiment. Only the seven steps from step two to step eight were used.

5.1.2. *Procedure.* The experiment was conducted in much the same way as experiment 1, using PsyScope 1.2.5 with on-screen instructions. For each subject, this experiment involved 84 trials, such that each of the seven continuum steps was heard twelve times, four times in isolation, and four times in each of the two different constructional contexts. No filler trials were given. Stimuli were presented in randomized order; the relative positions of the orthographical representations were kept constant.

5.2. *Results*

Figure 3 shows that syntactic context actually biases lexical identification in the predicted way. The diagram shows all *worse* responses relative to the three conditions of the experiment. The perceptual crossover covers all steps except the first one, regardless of context. It can be seen that syntactic context has an effect on the interpretation of the ambiguous stimulus, as the condition *it's getting* produces the most *worse* responses. The light grey curve, representing the condition *it's always* is flatter and drawn more to the left than the black curve, as this condition yields the fewest *worse* responses. The medium grey line, representing *worse* responses in the absence of a carrier phrase, falls between the two other lines, except at step 7.

A repeated measures ANOVA was conducted for all *worse* responses to measure the effect of the constructional carrier phrases. The calculation is based on all *worse* responses of fifteen subjects in three different conditions (isolation, it's always, it's getting) across the seven steps of the synthesized continuum. The constructional effect is significant in a by-subject analysis ($F_{(2,28)} = 13.60$, $p < 0.001$) and in the corresponding by-item analysis ($F_{(2,12)} = 18.83$, $p < 0.001$).

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 496)

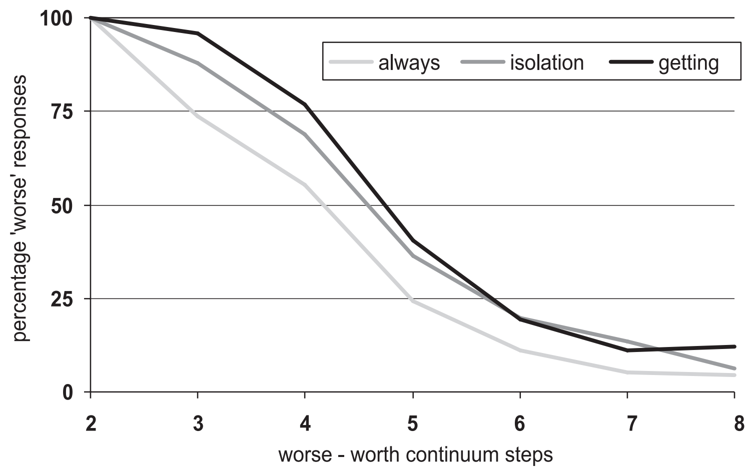

Figure 3. ‘worse’ *responses relative to condition*

6. Experiment 3—Compensation for coarticulation

The first two experiments yield evidence that knowledge of constructions and collocations induces shifts in phonemic category boundaries. This can be interpreted as a syntactic effect on phonological processing. While the explanation of such an effect requires some auxiliary assumptions on a modular view of grammar (Newmeyer 1998, 2003), the effect itself does not amount to a refutation of modularity. One possible criticism is that the observed results are the effect of late feedback between modules, which is how the effect of phonemic restoration due to semantic context (Warren and Warren 1970) is most appropriately interpreted. It could thus be that an input that is passed from the phonological module to the syntactic module is left unspecified or subsequently judged to be a misperception, and therefore re-analyzed at a relatively late processing stage. In order to show that syntactic effects on speech perception apply immediately at the level of auditory input processing, it needs to be demonstrated that on-line phonetic processing is affected by syntactic context. The third experiment investigates whether this is actually the case.

A potential source for such evidence is the phonetic effect of compensation for coarticulation (Elman and McClelland 1988). In processing naturally occurring speech, hearers do not expect each token of a phonemic category to be invariant. Hearers unconsciously compensate for the fact that every speech sound is influenced by its preceding elements. If a transition from one phoneme to the next takes effort, hearers will accommodate the resulting undershoot and perceive even less than perfect examples of a phoneme as a proper member of its category. The behavior of compensation for coarticulation can be exploited in an experimental setting. What the experiment aims to test is whether compensation for coarticulation, as an on-line phonetic effect, can be triggered by the syntactic context of a given construction.

To this end, not only one ambiguous stimulus is required, but two. For the first part of the complex ambiguous stimulus, the third experiment reuses stimuli of the second experiment. Subjects are presented with stimuli that are ambiguous between *worse* and *worth* in two conditions. In the control condition, subjects hear the stimulus in isolation, while the second condition presents the stimulus appended to the carrier phrase *it's always*. As has been shown in experiment 2, this carrier phrase biases lexical identification towards the competitor *worth*. It is assumed here that prior exposure to the stimuli does not have a biasing effect; the same subjects heard the enhanced stimuli in the third experiment.

The stimulus continues with an element that is phonetically ambiguous between the words *trying* and *crying*. It is here that compensation for coarticulation comes into play. The interpretation of the first stimulus (*worse*-*worth*) should lead subjects to categorize the second stimulus (*trying*-*crying*) in different ways, depending on the degree of effort on the part of the speaker to make the transition. Transitional effort is operationalized here, in a somewhat simplistic but not confounding way, as distance in production site. The two ambiguous stimuli yield four possible interpretations, which are shown in Table 3.

If the first stimulus is perceived as *worth*, which ends on the interdental /θ/, subjects should 'forgive' that a following velar /k/ is pronounced somewhat more towards the front; so they should be more likely to perceive the second stimulus as *crying*. Put simply, as /θ/ is produced further in the front of the mouth than /s/, we expect it to generate a relatively greater tolerance. By contrast, the transitions from dental to velar (/s/ > /k/), from interdental to alveolar (/θ/ > /t/) are relatively easy; and the transition from dental to alveolar (/s/ > /t/) is the easiest option al-

Table 3. *Stimuli interpretations and degree of effort in coarticulation*

Interpretation	Transition	Effort
worth crying	interdental - velar	difficult
worse crying	dental - velar	intermediate
worth trying	interdental - alveolar	indermediate
worse trying	dental - alveolar	easy

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 498)

together. Hearers should therefore be the least tolerant with respect to this transition.[2]

On the research hypothesis, the constructional context *it's always* should bias subjects towards perceiving *worth* more often than in the control condition. This, in turn, should result in a bias to perceive the second ambiguous stimulus as *crying* more often. If we thus observe more *crying* responses in the second condition, this would suggest that syntax affects even low-level phonetic processing.

6.1. *Method*

6.1.1. *Materials.* The carrier phrase *it's always* and the seven-step *worse*-*worth* continuum from the second experiment were re-used without further changes, but the stimuli were concatenated with further material. A ten-step continuum was created from the recorded items *crying* and *trying*, using the previously discussed method. Here, the first point and the last three points of the continuum were interpreted unambiguously in pilot tests, such that they were discarded and not used in the actual experiment. Only the six steps from step two to step seven were used. The subsequent concatenation of stimuli types yielded 42 different stimuli (seven *worth*-*worse* continuum steps times six *trying*-*crying* continuum steps).

6.1.2. *Procedure.* The experiment was conducted in the same way as the other experiments, using PsyScope 1.2.5 with on-screen instructions. The only difference concerned the fact that this time, subjects had to identify a percept as a word of English in a four-way choice: *worth crying*, *worse crying*, *worth trying*, or *worse trying*. Orthographical representations in a 44 pt font were either arrayed into the four corners of the screen or displayed to the left and right side of the screen. The same orthographical representation would appear in the same place throughout. For each subject, this experiment involved 252 trials. Each of the 42 stimuli was heard six times, three times in isolation, and three times after the carrier phrase *it's always*. No filler trials were given. Stimuli were presented in randomized order while the relative positions of the orthographical representations were kept constant.

2. A reviewer points out that both *worth* and *worse* should trigger the fronting of the initial consonant of *crying* and asks whether a minor difference in place of articulation, such as dental vs. interdental, has been shown to make a significant difference in compensation for coarticulation. Elman and McClelland (1988) report an effect for the alternation between /s/ and /ʃ/, i.e., an alveolar and a postalveolar fricative, so that indeed minute differences seem sufficient for the effect to obtain.

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 499)

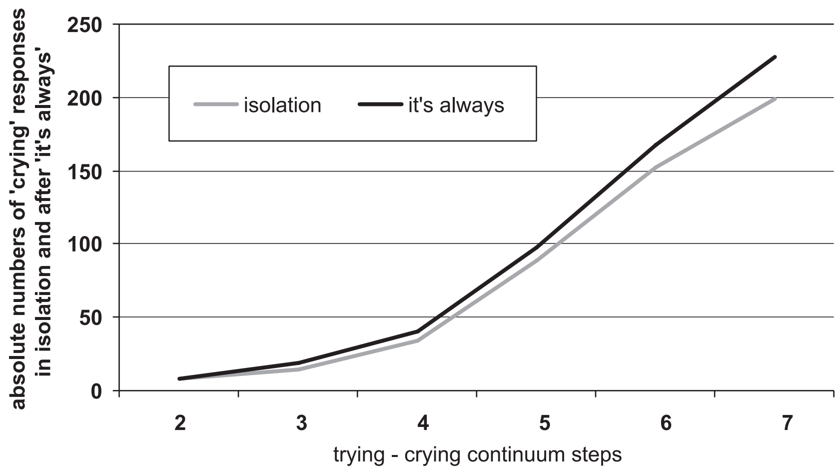

Figure 4. Crying *responses to the second ambiguous stimulus by condition*

6.2. *Results*

Figure 4 shows that there are fewer *crying* responses in the control condition than in the condition that involved the constructional carrier phrase. The diagram shows the absolute numbers of responses to the second ambiguous stimulus (*crying*-*trying*). After the constructional carrier phrase *it's always* we expected a higher number of *worth* responses and consequently a higher number of *crying* responses. This expectation is borne out.

A repeated measures ANOVA was conducted for all *crying* responses to measure the effect of the constructional carrier phrases. The calculation is based on all *crying* responses of fifteen subjects in two different conditions (isolation, it's always) across the six steps of the synthesized continuum. The constructional effect approaches significance in a by-subject analysis $F_{(1,14)} = 4.02$, $p = 0.065$); the corresponding by-item analysis returns a significant result ($F_{(1,5)} = 6.58$, $p = 0.050$).

7. Conclusion

The results of the three experiments demonstrate that syntactic context in the form of constructions and collocations has an effect on both phonemic categorization and low-level phonetic processing. Presenting ambiguous sounds in the carrier phrase of a constructional or collocational frame alters the phonemic category boundaries in a lexical identification task, and it can induce the phonetic effect of compensation for coarticulation. It needs to be acknowledged that the observed effects fail to reach significance in the cases of the by-item analysis of experiment 1 and the by-subject analysis of experiment 3. Here, there is only evidence in the form of trends. The directions of the observed effects are, however, as predicted; they always move towards the lexical element that more frequently occurs with the carrier phrase. This reaffirms the point that collocations and collocational patterns within constructions (Stefanowitsch and Gries 2003) have a psychological reality that shapes the way in which hearers perceive speech. It can also be concluded that the lexically based Ganong-effect has a more abstract counterpart which extends to the level of syntax, and which is not restricted to the opposition of words and non-words. The result that subjects are biased towards hearing entrenched units over hearing chance collocations is consistent with views held in Construction Grammar and cognitive linguistics, but up to now, this view had not been sufficiently supported through empirical studies. The results of the present study provide new evidence that syntactic and lexical knowledge are not stored in different mental modules, but rather form a continuum from heavily entrenched and conventionalized units to loosely connected elements (Bybee 2005).

Received 27 June 2007
Revision received 3 January 2008
ICSI Berkeley, USA

References

Barlow, Michael and Suzanne E. Kemmer
2000 *Usage-based Models of Language*. Stanford: CSLI.

Boersma, Paul and David Weenink
2005 Praat: doing phonetics by computer (Version 4.3.14) [Computer program]. Retrieved May 26, 2005, from http://www.praat.org/.

Bybee, Joan L.
2000 The phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow and S. Kemmer. (eds.), *Usage-based Models of Language*. Stanford: CSLI.
2001 *Phonology and Language Use*. Cambridge: Cambridge University Press.
2006 From usage to grammar: The mind's response to repetition. *Language* 82(4), 711–733.

Bybee, Joan L. and Paul Hopper (eds.)
2001 *Frequency and the Emergence of Linguistic structure*. Amsterdam: John Benjamins.

Clifton, Charles Jr.
1991 Syntactic modularity in sentence comprehension. In R. R. Hoffman and D. S. Palermo (eds.), *Cognition and the symbolic processes, Vol. 3: Applied and ecological perspectives*. Hillsdale, NJ: Erlbaum, 95–114.

Cohen, Jonathan, Brian MacWhinney, Matthew Flatt, and Jefferson Provost
1993 PsyScope: An interactive graphical system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers* 25, 257–271.

Cooper, William E. and Jeanne Paccia-Cooper
1980 *Syntax and Speech*. Cambridge, MA: Harvard University Press.

Elman, Jeff L. and James L. McClelland
1988 Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language* 27, 143–165.

Fillmore, Charles J., Paul Kay, and Mary C. O'Connor
1988 Regularity and idiomaticity in grammatical constructions: The case of Let Alone. *Language* 64(3), 510–538.

Fodor, Jerry A.
1980 *The Modularity of Mind*. Cambridge: Bradford books.

Francis, W. Nelson and Henry Kucera
1982 *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton-Mifflin.

Frazier, Lyn
1987 Sentence Processing: A tutorial review. In M. Coltheart (ed.), *Attention and Performance XII: The Psychology of Reading*. Hove: Erlbaum, 559–586.

Fry, Dennis B., Arthur Abramson, Peter D. Eimas, and Alvin M. Liberman
1962 The identification and discrimination of synthetic vowels. *Language and Speech* 5, 171–189.

Gahl, Susanne and Susan M. Garnsey
2004 Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80(4), 748–775.

Ganong,William F.
1980 Phonetic categorization in auditory perception. *Journal of Experimental Psychology: Human Perception and Performance* 6, 110–125.

Goldberg, Adele E.
1995 *Constructions*. Chicago: University of Chicago Press.
2006 *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.

Gregory, Michelle, W. D. Raymond, A. Bell, E. Fosler-Lussier, and D. Jurafsky
1999 The effects of collocational strength and contextual predictability in lexical production. *Chicago Linguistics Society* 35.

Hooper, Joan B.
1976 *An Introduction to Natural Generative Phonology*. New York: Academic Press.

Jurafsky, Dan, Alan Bell, Michelle Gregory, and William D. Raymond
2001 Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In J. L. Bybee and P. Hopper. (eds.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, 229–254.

Kemmer, Suzanne E.
2001 *Causative Constructions and Cognitive Models: The English Make Causative*. The First Seoul International Conference on Discourse and Cognitive Linguistics: Perspectives for the 21st Century, 803–832.

Kewley-Port, Diane and Paul A. Luce
1984 Time-varying features of initial stop consonants in auditory running spectra: A first report. *Perception and Psychophysics* 35, 353–360.

Langacker, Ronald W.
1987 *Foundations of Cognitive Grammar*. Stanford: Stanford University Press.
2005 Construction grammars: Cognitive, radical, and less so. In F. J. Ruiz de Mendoza Ibáñez and M. S. Peña Cervel (eds.), *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interaction*. Berlin: Mouton de Gruyter, 101–159.
Leech, Geoffrey
1992 100 million words of English: the British National Corpus. *Language Research* 28(1), 1–13.
Liberman, Alvin M., Katherine S. Harris, H. S. Hoffman, and B. C. Griffith
1957 The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54, 358–68.
Magnuson, James, Bob McMurray, Michael Tanenhaus and Richard Aslin
2003 Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science* 27(2), 285–298.
McClelland, James L., David E. Rumelhart and the PDP Research Group
1986 *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*. Cambridge, MA: MIT Press.
Newmeyer, Frederick J.
1998 *Language Form and Language Function*. Cambridge: MIT Press.
2003 Grammar is grammar and usage is usage. *Language* 79, 682–707.
Stefanowitsch, Anatol and Stefan Th. Gries
2003 Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8, 209–43.
Tanenhaus, Michael, Michael Spivey-Knowlton, Kathleen Eberhard and Julie Sedivy
1995 Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 1632–1634.
Warren, Richard M. and Roslyn P. Warren
1970 Auditory illusions and confusions. *Scientific American* 223, 30–36.
Zipf, George K.
1935 *The Psycho-biology of Language*. Boston: Houghton-Mifflin.

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 503)

(AutoPDF V7 27/5/08 11:48) WDG (148×225mm) TimesM J-1972 COGL, 19:3 PMU: WSL(W) 23/05/2008 pp. 483–504 1972_19-3_06 (p. 504)