Adaptation Effects in Lexical Processing

Psycholinguistic research generally adopts a scientific strategy that assumes a relatively stable set of representations and processes. In accordance with this strategy, researchers average measurements across trials, in an attempt to get a statistically stable estimate of performance for a given experimental condition. In this paper, we present four sets of example data drawn from various psycholinguistic tasks and show that the psycholinguistic system appears to adapt across the trials of the experiments. We show that there are cases in which a factor has no main effect, but interacts across trial; in other cases there is a main effect of a factor, but that factor also interacts with trial. Finally, we show that there are some cases in which the way that a factor interacts across trials is dependent on other, unrelated conditions included in the experiment. Our discussion focuses on both theoretical and methodological implications of the adaptiveness of the psycholinguistic system.

1. Introduction

Psycholinguistic research is concerned with how language is represented and used. An important area within this field is lexical processing, which refers to the means through which words are accessed. Lexical processing has received considerable attention because the notion that word access precedes other aspects of language processing, such as the construction of phrases and sentences, is central to most psycholinguistic theories. For example, the general idea (i.e., the central assumption behind most theories of word processing) is that words (or parts of words) are stored in a mental lexicon and that various representations become activated when someone reads or hears words before a single candidate representation is chosen and remains active. In this sense, lexical processing involves the recognition of lexical units and these units are then used as the basis for other linguistic processing (including sentence comprehension).
To discover how the lexical system is structured and accessed, researchers use various experimental techniques including lexical decision, naming, and monitoring eye-movements during reading. A dominant methodological assumption in psycholinguistic research (and in experimental psychology research, in general) is that an experiment provides a snapshot of the processing system. That is, the data provide insight into the structure of the mental representations and associated processes that allow the person to perform a particular task. Thus, when analyzing the data, researchers often average data across trials to provide more stable estimates (Ratcliff 1979; Whelan 2008). In this paper, we discuss whether the assumption of a stable system is always appropriate and provide four examples taken from recent experiments that provide evidence of rapid adaptation in lexical processing tasks.

1.1 The Mental Lexicon as a relatively stable set of representations

An unresolved issue for psycholinguistic research concerns the nature and access of lexical representations. In particular, researchers have been trying to determine whether complex words are represented and accessed as whole units or in terms of smaller units, such as their constituent morphemes. A guiding principle behind the architectures proposed for the processing of complex words has been the balance of lexical storage vs. morphological computation (see Libben 2010; Sandra 1994 for a discussion of this issue) and psycholinguistic theories of word processing vary in terms of how much emphasis is placed on the storage of whole-word representations and how much emphasis is placed on computation. Some theories posit that all words are stored as whole-word forms (Lukatela, Carello, & Turvey 1987; Manelis & Tharp 1977). Other theories posit that words are represented and accessed via their constituent morphemes, which are later composed to form the whole-word (Chialant & Caramazza 1995; Dell 1986; Frauenfelder & Schreuder 1991; Laudanna & Burani 1985; Schreuder & Baayen 1995; Taft & Forster 1975, 1976). The latter theories vary in terms of the point at which the constituents’ representations become available (see Kuperman, Bertram, & Baayen 2010 for an overview).

Despite their differences in terms of assumptions concerning the nature of lexical representations, the theories are similar in that they assume that the mental lexicon is relatively stable and that changes occur over longer periods of time as new words are added to the system and the representations of infrequent words are weakened. That is, once added to the mental lexicon, the representation for a word is not changed beyond the activation of the particular representation. For example, according to Schreuder and Baayen’s (1995) framework, if a novel complex form is encountered, the system computes a meaning and leaves a memory trace in the form of a concept node connected to an access representation. This representation decays and, thus, the word must be re-encountered to be maintained in the mental lexicon. Once added to the mental lexicon, the nature of the representation is not greatly altered, although the strength of the representation might change.
Related to this point, the way in which the system processes words is not expected to change within a short timeframe. In dual-route models, for example, morphologically complex words can be accessed either via their whole-word representation (e.g., blueberry) or via the constituent representations (e.g., blue and berry). Which route is most successful depends on a variety of factors including the frequency of the word (Schreuder & Baayen 1995) and the relative frequency of the constituent representations and the whole-word representation (e.g., Colé, Segui & Taft 1997; Hay 2001). However, the main point is that the relative success of the whole-word route versus constituent-access route is based on the properties of a particular word. That is, emphasizing the constituents of blueberry and thereby aiding the decomposition route does not subsequently alter the relative success of these two routes for an unrelated word such as notebook. We will return to the importance of this point in Section 3.

On the face of it, the assumption that the Mental Lexicon is a stable system is not unreasonable if researchers are attempting to model adult language users. Thus, relatively little attention has been paid to whether or how the language system adapts, especially within the short timeframe of a single experiment. Although there has been some work conducted, for example, on the way in which processing is affected by the nature of the task (e.g., Duñabeitia, Kinoshita, Carreiras and Norris 2011; Paterson, Alcock, & Liversedge 2011) or the materials (e.g., Keuleers, Diependaele & Brysbaert 2010), much more research has been focused on the nature of the representations that allow people to identify and use words in language processing. That said, the question of storage versus computation is becoming more prevalent (Libben 2005).

However, the assumption that the access of lexical representations remains constant throughout the experiment might be problematic. As we discuss below, there is some reason to believe that the language system in general is quite adaptable and may rapidly tune itself to the current processing context. Examining adaptation may provide useful insight into the language system because it would allow researchers to tease apart assumptions about representation from assumptions about processing. It does so by allowing researchers to identify which aspects of the system are relatively stable from one experimental setting to another and which aspects are able to adjust quickly.

1.2 Sensitivity and adaptability of language system

The psycholinguistic literature contains many examples that illustrate that lexical processing is sensitive to the current experimental context and that the system rapidly adapts. In this section, we provide a brief overview of a few of these examples.

Research using a priming paradigm illustrates that the language system is sensitive to recent experience as well as to long-term associations among various sources of linguistic information (such as orthography, phonology, morphology, and semantics). When using this experimental paradigm, researchers examine whether exposure to a word (called a prime) alters subsequent pro-
cessing of another word (called the target). The theoretical assumption is that the processing of the prime alters the state of the mental lexicon by altering the activation of various representations. This change in activation affects the subsequent processing of related representations.

Numerous studies using a priming paradigm have found that lexical processing is affected by exposure to a related prime. First, the system is sensitive to orthography. The repetition of letters affects ease of processing; presentation of the nonword bontrast speeds the processing of contrast (Davis & Lupker 2006; Forster & Veres 1998). Second, the system is sensitive to phonology. For example, Slowiaczek and Hamburger (1992) found that it takes less time to recognize a word (e.g., dale) when preceded by a word that has similar sound (e.g., tale) than when preceded by a phonologically unrelated word (e.g., book). Third, the system is sensitive to morphology. Recent exposure to a compound word facilitates the subsequent processing of its constituents (Masson & MacLeod 1992; Sandra 1990; Weldon 1991; Whittlesea & Brooks 1988). Weldon (1991) found that the identification of the target word black was faster when preceded by either blackbird or blackmail. Likewise, Masson and MacLeod (1992) demonstrated that a constituent (e.g., break) is more accurately identified when it has been seen as part of a noun phase (e.g., coffee break) than when it has not been previously studied. Constituent priming occurs even when the prime is not an existing word. For example, both copper block (a novel phrase) and sympathy block (a nonsensical item) facilitate the processing of block (Osgood & Hoosain 1974). Fourth, the system is sensitive to semantics. It takes less time to recognize the word doctor when it is preceded by a semantically related word, such as nurse, than when it is preceded by an unrelated word, such as grass (Meyer & Schvaneveldt 1971).

In addition to being sensitive to recently processed items, lexical processing is affected by the nature of nonword items in lexical decisions tasks. For example, the use of nonwords with transposed letters (jugde) increased response time for both words and nonwords (Perea & Lupker 2004). Differences in word length (i.e., number of letters) also affects processing; Chumbley and Balota (1984) found faster responses to the experimental items in an experiment in which the nonwords were on average one letter shorter than the words, than in an experiment where the nonwords and words were matched for average letter length. Other research (Rastle & Brysbaert 2006) has found that lexical decision responses are faster when the preceding prime is a pseudohomophone. The degree to which the nonwords mimic real–words also affects lexical processing (Keuleers, Diependaele, & Brysbaert 2010).

Another example illustrating the adaptability of lexical processing comes from research on referential communication and indicates that speakers rapidly adjust their speech to suit the current situation. Conversational partners often converge on an expression and will persist in using that expression even when there is no longer a need to include the additional information. To use an example from Brennan and Clark (1996), the term pennyloafer was initially used during a conversation to denote a particular shoe among other possible shoes. However, the speaker continued to use this term even when no other
shoes were present in the display. More recent work has suggested that over-specification might be beneficial to communication in that it speeds up identification processes (e.g., Arts, Maes, Noordman, & Jansen 2010). In addition, modifying information is used to shift the addressee’s focus of attention (Ariel 1990; Chafe 1994; Gundel, Hedberg, and Zacharski 1993; Prince 1992).

In sum, there is good reason to believe that the language system adapts as it is exposed to different processing demands. Some of this adaptation is tightly tied to particular target items (e.g., repetition priming effects), but some appears to be a wider form of adaptation, such as the effects of different kinds of filler items. Adaptation tied to particular items may not cause us much concern about using the “experimental snapshot” metaphor that we described above, but wider forms of adaptation do raise concerns that our experimental techniques might not provide the kind of clean and direct look at lexical representations that are usually assumed.

1.3 Overview of current investigation

Our aim is to examine whether and how quickly the lexical system adapts to processing demands. Thus, we will present four examples to examine questions about whether the language processing system adapts within the timeframe of a single experiment, to examine whether the changes are item-specific, and to demonstrate that considering adaptation can lead to a different conclusions about the nature of lexical processing and the mental lexicon relative to when adaptation is not considered.

In terms of statistical analysis, linear mixed-effects (LME) analysis is well-suited for investigating these issues. LME regression models include both fixed effects (such as experimental variables) and random effects (such as subjects and items) and can be conducted in a variety of statistical software packages including S and S-Plus (Pinheiro & Bates 2000), R (Baayen 2008), and Stata (Rabe-Hesketh & Skrondal 2012). LME models are especially useful for repeated measures designs (i.e., designs in which all conditions are presented to the participants); Baayen, Davidson, and Bates (2008) demonstrate that this procedure has several advantages over ANOVA analyses. Two particularly relevant advantages for our current investigation is that LME analysis does not require data to be aggregated, and allows for the inclusion of continuous variables as the primary variables of interest. These two properties allow us to examine the influence of trial and, more importantly given our focus on adaptability, whether the influences of our independent variables change across trials. That is, we can use the technique to determine whether trial interacts with the independent variables.

In all of the analyses reported in the subsequent sections, we included items and subjects as crossed random effects which is an effective way of solving the “language—as—a—fixed—effect fallacy” problem that was identified by Clark (1973, see also Baayen, Davidson, & Bates 2008; Raaijmakers, Schrijnemakers, & Gremmen 1999). Our analyses were conducted using the xtmixed function in STATA (Statacorp 2013). We report tests conducted on the esti-
mates of the fixed effects. These statistical tests are based on the null hypothesis that the slope is zero and, thus, the tests evaluate whether the fixed effect is a valid predictor of the dependent variable (i.e., whether the independent variable influences the dependent variable).

2. Example Datasets and analyses

In this section, we present several examples drawn from recently conducted experiments that each illustrate the usefulness of considering adaptation across trials. In each case, the conclusions drawn from an analysis that does not take into account adaptation differ from the conclusions drawn from an analysis that does take adaptation into account. The experiments were originally conducted to test different types of hypotheses about the nature of compound word processing and are selecting from several research projects. For the purpose of this paper, we focus instead on subsets of the data from those experiments to illustrate particular points about adaptation.

2.1 Example 1

In this experiment, we examined whether lexical decision latencies for compound words were affected by the semantic transparency of each of the two constituents. Semantically transparent constituents (e.g., moon in moonlight) contribute to the compound’s meaning, whereas semantically opaque constituents (e.g., comb in honeycomb, chop in chopstick, and hog and wash in hogwash) do not. In the experiment, we crossed transparency of the first constituent with transparency of the second constituent to yield four experimental conditions: transparent–transparent (e.g., paintbrush), transparent–opaque (e.g., gingersnap), opaque–transparent (e.g., bulldog) and opaque–opaque (e.g., buttercup). For the purpose of this example, we will focus on the question of whether the transparency of the second constituent (which we will refer to as “C2 transparency”) influences the time to indicate that the compound is a word. Lexical decision times were log transformed to reduce skewness because the model assumes that the dependent variable is normally distributed.

A LME model including C2 transparency as a predictor variable indicated that this variable did not influence lexical decision times, $z = 1.42, p = .16$. Based on this result, one would conclude that C2 transparency does not influence lexical decision time and that C2 transparency is either not represented in the mental lexicon or that it is not involved in the access of the compound’s whole-word representation.

However, it is possible that the influence of this variable might be obscured by variability in the data. Generally, during an experiment, people speed up as they become more practiced at the task and this can contribute to variation in response times. Thus, a second model included trial number as a predictor variable to reduce the unexplained variability in the model. Although this second analysis indicates an influence of trial, $z = -10.19, p < .0001$, there is still no indication that C2 transparency affects lexical decision time, $z = 1.26$, $p = .21$. 

132
Although this second model takes into account variability due to trial, it does not examine whether the system is adapting across the experiment. To examine this possibility, we fit a model that includes the interaction between trial and C2 transparency and we found that, indeed, this interaction is statistically significant, $z = 2.87, p = .004$. The model with the interaction fits the data better than a model that does not include the interaction, $\text{Chi} = 8.23, p = .004$. The nature of this interaction is illustrated in Figure 1.

This result indicates that C2 transparency does affect lexical decision times, but that participants are adapting across trials. Importantly, the changes in performance that were occurring across the experiment were not “across the board” changes in which participants were speeding up as they gain experience with the task. Instead, the adaptation differentially affected compounds with opaque and transparent heads (i.e., second constituents). It could be the case that people are getting better at suppressing information about the second constituent (i.e., the head of the compound). This suppression helps the processing of compounds with opaque heads, but removes the advantage for transparent heads that is seen earlier in the experiment. It is also worth noting that the items were never repeated and thus the adaptation is to a general type of item (i.e., to opaque heads and to transparent heads) rather than to specific items.

Figure 1: Cross–trial effects of the semantic transparency of the second constituent on lexical decision time
2.2 Example 2

One might think that the adaptation effect we found in the previous experiment is due simply to the fact that lexical decision is a rather unusual and arguably unnatural task—after all, in real life we are rarely called upon to explicitly determine whether a letter string is or is not a word. In this example, we focus on adaptation in a task that is more practiced and natural (especially for undergraduate students)—namely, written production as measured by typing times (Libben, Weber, & Miwa 2012; Sahel, Nottbusch, & Grimm 2008). In the experiment, we used a priming paradigm in which a compound was preceded by a word that was either the same as the first constituent of the compound or different. For example, staircase was preceded by either stair or liner. The compounds always had opaque heads, but varied in terms of the transparency of the first constituent. The prime (e.g., stair or liner) was presented using a masked priming paradigm (Forster & Davis 1984) in which the prime was briefly presented on a computer screen and was followed by a series of hash-marks (#####) and finally by the compound. Participants typed in the compound using the keyboard as the computer recorded the time required to type each letter. For the purpose of this example, we focus only on the influence of prime on the typing time for the first letter of the second constituent. This measure provides information about whether participants are sensitive to morphemic structure. The times were log transformed to reduce skewness.

A simple model including only prime indicates that prime does not affect the typing latency, $z = .28, p = .78$. That is, making the first constituent of the compound more available (via the related prime) did not appear to affect the production of the second constituent. A second model including trial showed that typing latency gets faster across trials, $z = -2.44, p = .02$, but that prime did not exert an influence, $z = .30, p = .77$. If a researcher only considered these two models, then he/she would conclude that increasing the availability of the first constituent did not affect the ease of producing the second constituent.
However, if we consider the possibility that the influence of the prime is changing across the experiment, then we see that activating the first constituent does affect production of the second constituent (see Figure 2). A model including the interaction term revealed a prime by trial interaction, $z = -2.75$, $p = .006$; the model with the interaction term provided a better fit to the typing data than did a model without this interaction, $\text{Chi} = 7.53$, $p = .006$. This analysis indicates that enhancing the first constituent (via the presentation of the prime, e.g., *stair*, prior to *staircase*) initially slowed typing of the first letter of the second constituent, but then, later in the experiment, the processing system took advantage of this cue, perhaps by boosting activation to the entire morphemic structure, which speeded the time required to initiate the second constituent. However, this adaptation appears to cause interference from the different prime (e.g., *liner* prior to *staircase*) because there is no speed-up across trials for this condition.

### 2.3 Example 3

In this experiment, we again used typing time of the first letter of the second constituent as our dependent variable. The times were log transformed to reduce skewness. Using a masked priming paradigm, a compound (e.g., *popsyseed*) was preceded by the brief presentation of word that was either semantically related (e.g., *opium*) or unrelated (e.g., *arcade*) to the compound’s first
constituent. The compounds had semantically transparent heads, but varied in semantic transparency of the first constituent in that half the compounds had transparent first constituents and half had opaque first constituents.

The simple model indicates that typing is faster when the prime was related than when the prime was unrelated, $z = -2.30$, $p = .02$. A second model that includes trial as a predictor variable shows an effect of the prime, $z = -2.30$, $p = .02$, but no effect of trial, $z = -1.13$, $p = .26$. Unlike the examples in Sections 2.1 and 2.2, the analysis without taking into account adaptation indicates that the presentation of a semantically related prime affects lexical processing.

However, a model that includes an interaction term shows that the effect of prime is not constant across trials, $z = 2.08$, $p = .04$; the model with the interaction term fits the data better than does the model without the interaction term, $Chi = 4.32$, $p = .04$. As shown in Figure 3, there is an advantage of the related prime condition relative to the unrelated prime condition early in the experiment but not in the later part of the experiment. The analysis shows that participants were adapting across trials such that the time to type the first letter of the second constituent (e.g., the $s$ in *poppyseed*) gets faster during the experiment when the compound was preceded by an unrelated prime, but slows down when preceded by a related prime.

![Figure 3: Cross-trial effects of prime on typing time of the first letter of the second constituent](image-url)
2.4 Example 4

The previous examples show that the lexical processing system adapts across the course of the experiment. In this section, we examine whether the degree of adaptation is affected by which experimental conditions are included. This question directly focuses on how stable the adaptation is across experiments and, in particular, whether adaptation is based on each experimental condition regardless of which conditions are included in the particular experiment, or whether adaptation is dependent on the particular combination of experimental conditions.

Thus far, we have shown evidence of adaptation over trials, and that adaptation is not just getting better across the board, nor is it the same for all conditions. These results strongly suggest that the system is adapting to the whole task in a flexible way. If the language system is adapting to the processing demands of the whole task, then we would expect that the adaptation would be sensitive not only to the type of task (e.g., lexical decision, typing, etc.), but also to the particular experimental conditions. Consequently, processing changes (or lack of change) across trials for a particular condition might change across experiments if conditions are added or excluded from the experiment.

To examine this prediction, in this section, we compare two experiments that examine the influence of a word containing the same initial letters on the processing of a subsequent word. In the first experiment, we used a masked priming paradigm (Forster & Davis 1984) to examine whether lexical decision times to the target (e.g., car) are differentially affected when either a pseudocompound (e.g., carrot, which is monomorphemic but contains two free morphemes, car and rot and, thus, has the appearance of a compound) or an orthographically-related word (e.g., career) is the masked prime. The orthographically-related words began with the same embedded word (e.g., car) as the pseudocompounds, but the remaining letters (e.g., eer) did not correspond to a morpheme and, consequently, these words did not appear to have a compound-like structure. In the second experiment, we include an unrelated prime condition (e.g., mutant) as well. In the analyses, we examine whether the log transformed response time to the target (e.g., car) is predicted by prime-type. If so, then this indicates that exposure to the prime word alters the activation of the representation of the target word.

In the first experiment, responses got faster across trials ($z = 1.94, p = .05$), but this decrease in response time was equivalent for the pseudocompound and orthographic conditions; prime and trial did not interact, $z = .64, p = .52$. For example, the time to correctly respond “word” to car was the same when the prime carrot was briefly presented as when the prime career was presented and people got faster at making lexical decision judgments as the experiment progressed (see Figure 4).
In the second experiment, which included an unrelated condition, a prime by trial interaction emerged, *Chi* = 6.18, *p* = .05, and the nature of this interaction is shown in Figure 5. A comparison of Figures 4 and 5 clearly shows that the inclusion of the unrelated condition in the second experiment changed the nature of the adaptation for the pseudocompound and orthographic conditions. Across the experiment, lexical decision times to the target word (e.g., *car*) became faster when the masked prime was orthographically related (e.g., *career*) but become slower when the masked prime was a pseudocompound (e.g., *carrot*). That is, presentation of a pseudocompound made it more difficult to judge that the target word was indeed a word and this difficulty increased during the experiment. The presentation of an orthographically related word initially made it more difficult to respond to the target word, but the difficulty decreased during the experiment.
Taken together, the analysis suggests that the system was attempting to reduce the difficulty caused by the presentation of a prime that contained the word to be judged but that the pseudocompounds and orthographically related words produced different processing demands and could not both be offset by the same adjustments to the system. This suggests that the system is differentially sensitive to pseudo-morphemes and orthographic strings.

3. Theoretical Implications

These data sets provide four examples that suggest that the language system can rapidly adapt within the timeframe of an experiment. Furthermore, it is important to note that these studies only took approximately 15–20 minutes to complete. Thus, one conclusion that should be drawn from these examples is that lexical processing is, in some cases, highly and rapidly adaptive. This, in itself, is an important characteristic of the language system, and psycholinguistic theories must accommodate this fact.

In addition, it is worth pointing out that the forms of adaptation in our examples are not all particularly simple ones. For example, one might get an interaction across trial even if all that was really happening was that the total performance was getting faster. For example, when performance is slow overall, it might be easier to detect a difference between Condition A and Condition B, but it might be hard to detect such a difference when overall
performance is fast. None of the examples we provide above has this simple character. Nor do they have the character of simply improving performance in all conditions, but perhaps just a bit faster in one condition than in another (as would be the case if the improvement were purely attributable to a practice effect). Of course, this does not mean that such cases of adaptation do not occur! Indeed, one might hope that the language system is such that in many cases it would adapt so that all conditions get better with experience, and that the worst performing conditions improve the most. In some sense, that would be the optimal kind of adaptation to see. Our examples, however, suggest that there are combinations of tasks and conditions for which the system simply cannot adapt in a way that is optimal across all tasks and conditions, or at least that it cannot adapt optimally without much more experience. Of course, to have an adequate theory of language processing will require an adequate theory of the system’s ability to adapt. Furthermore, the system’s ability to adapt, and the patterns of adaptation that are possible, also provide further constraints on how the system operates.

Finally, the data discussed in Section 2 suggest that we should be cautious about attributing effects solely to the architecture of the system. That is, assuming that the presence or absence of an effect directly reflects the representations in the mental lexicon and the links among them may not be justified. This assumption is problematic because such an approach does not take into account the processes that might be involved. For example, Ji, Gagné, and Spalding (2011) found that the processing of opaque compounds was slower than lemma–frequency–matched transparent compounds and that manipulations that aided morphological decomposition, such as presenting the first and second constituents in two different colours, slowed the processing of opaque compounds (e.g., hogwash), but not transparent compounds (e.g., snowball). These manipulations made it easier to identify the constituent morphemes which aided the meaning construction process. However, the meaning that is constructed based on the constituents is inconsistent with the established meaning of an opaque compound and, therefore, produces competition. These results indicate that there might be both facilitatory and inhibitory processes occurring and that, in some cases, these processes might offset each other such that no effect is observed. Thus, ascribing a lack of an effect to the absence of links among various representations (e.g., between the lexical and semantic representations of opaque constituents) is not always the most accurate way of interpreting the results. Furthermore, our Example 4 suggests that any attempt to build a representational explanation of the data in the pseudo-compound prime and orthographically related prime conditions will be unsuccessful, as the processing differs depending on the presence or absence of an unrelated baseline condition. Consequently, when constructing theories, it is useful to consider “what can the system do” rather than “what is the system’s structure”. In other words, both representation and processing must be taken into account (see Libben (2005) for a similar argument).

Libben (2010) also discusses how inhibition and facilitation might arise from morphemic representations. He argues that morphemes are position–
bound and, consequently, the representation of a morpheme used in the first position is not the same as the representation of a morpheme used in the second position, nor as the free morpheme. For example, *boardroom* contains a different morpheme (i.e., *board–*) than does *keyboard* (i.e., *–board*) and the word *board*. Libben posits that these representations might conflict with each other. Thus, if the word *board* is used as a prime for the target *keyboard*, then one would expect facilitation from the word–initial bound morphemic representation (i.e., *–board*) but inhibition from the free morpheme (i.e., *board*) and from the word–initial position morpheme (i.e., *board–*). Consistent with the prediction that various morphemic representations compete, Baayen (2010) finds processing costs arising from a morpheme’s use as both a modifier and head noun.

As these examples illustrate, lexical processing is affected by the particular representations that are activated and by the pattern of facilitation and inhibition arising from these representations. However, our examples in Section 2 and the findings by Ji et al. (2011) discussed in this section suggest that processing can be affected on a system–wide basis, as well. To illustrate, in Example 4 (Section 2.4), processing in the pseudocompound condition became more difficult as the experiment progresses even though the particular words or constituents were never repeated during the experiment. Thus, it is not the case that the specific lexical representations are being altered (i.e., the effects are not item–specific), but rather it is the way in which particular types of words (e.g., pseudocompounds) are being processed that is being affected. That is, adaptation is happening on a more general, abstract, level. This adaptation appears to affect how the representations are being used and also the relative weight that is being placed on the various types of representations, such as whole–word representations versus constituent representations. This finding suggests that the system is sensitive to commonalities among words of particular types (e.g., pseudocompounds, opaque compounds, transparent compounds, and monomorphic words). In this respect, data demonstrating differences in adaptation for different types of words provides useful insight into distinctions that are relevant to the processing system.

One might wonder whether changes in performance reflect adaptation or whether they arise from strategies or other types of effects. We argue that the patterns of data observed in our examples are adaptation, rather than strategy, for several reasons. First, strategic processing requires conscious awareness of the type of the items. Thus, to develop a strategy to respond differently in the pseudocompound and non–compound conditions, for example, the participant must be aware of the masked prime and then decide to use that knowledge in making a response to the target word. This possibility is unlikely because participants are not consciously aware of the identify of masked primes (Bodner & Masson 2004; Forster & Davis 1984). Also, it seems unlikely that participants would create a strategy to delay their responses to *car* following *carpet* because generally strategies are used to facilitate performance. Second, it is unlikely that participants have conscious control over lexical processing because this processing is highly practiced and nearly automatic as can be seen in Stroop
effect experiments (Stroop 1935) in which people must name the colour of the ink in which a word was displayed. In these experiments, participants cannot help but identify words even when doing so leads to slower performance as is the case when having to say “red” when the word blue is displayed in red. Third, the changes in performance happened very rapidly (i.e., after a few trials) and it is highly unlikely that participants could have identified a response strategy with so little experience. Finally, the pattern of the changes is not obviously related to any simple strategy. That is, it is not the case that the participants are simply deciding to respond “word” to the target if the prime is a word, or are deciding to respond “yes” if letters are repeated.

As alluded to earlier, the results are not attributable to practice effects. Practice effects are effects that occur across the course of an experiment as participants become more proficient at the task. However, practice effects are general across item type and, thus, one would not expect to see an interaction between trial and experimental condition. For the same reason, the results are not attributable to fatigue or lack of interest because there should be a general decrease in performance as the experiment progresses. The data do not show an overall speed up or drop off but rather different patterns of performance across trials based on condition.

The pattern of results also cannot be reduced to episodic effects. Episodic effects occur when the system recruits specific episodic memories. Bodner and Masson (2003; Masson & Bodner 2003) have argued that priming might be a type of episodic memory effect (cf. Kinoshita, Forster, & Mozer 2008; Kinoshita, Mozer, & Forster 2011). They argue that the prime forms an episodic resource that can be subsequently recruited. However, our results point to a different (or perhaps additional) source of influence. The order of items was randomized for each participant, and, consequently, item–specific episodic effects would be removed by this randomization. Moreover, each item was seen only once and, thus, the effect is not due to the retrieval of specific episodic traces. Finally, the notion of prime validity as a modulator of semantic priming (see Bodner & Masson 2003, 2004) does not apply in the empirical examples discussed in Section 2 because we were not manipulating the proportion of related–prime trials; the proportion did not change across the course of the experiment, nor does the proportion change across the experimental conditions.

Although the literature contains several examples of list effects (e.g., Feldman & Basnight–Brown 2008), the observation that the processing of the experimental conditions differentially changed across trial is not explained by this type of effect. List effects refer to processing differences brought about by different types of lists. However, within a given study, the list (i.e., the proportion of items in each condition) was the same. The difference between the two experiments in Example 4 is an example of a list effect, but only trivially so: The two kinds of lists (which, in this case, included different conditions, rather than the usual case of lists simply containing different items) gave different results, but the point of interest is the patterns of adaptation over trials, so that cannot be attributed to the different lists. That is, the list difference explains neither of the patterns of adaptation that are of interest.
In summary, then, the examples described above provide strong evidence that the lexical processing system is adapting over a relatively short period of time. These adaptation effects are specific to types of items, resulting in interactions of experimental conditions across trial. The adaptation effects are not reducible to other kinds of effects, such as simple practice effects, fatigue, or boredom, nor to item specific influences as seen in episodic or list effects.

4. Methodological implications

Examining the adaptability of the language and cognitive systems alters the scientific approach that is used because the data analysis needs to be able to evaluate changes occurring during the experiment. Clearly, in our examples, the conclusions that one would draw from the data are strongly dependent on whether or not the possibility of adaptation is examined in the analysis. Many of the traditional forms of analysis, including ANOVA and t–tests, involve summarizing data across trials, which obscures cross–trial effects. As our four examples illustrate, summarizing across trials can, in some cases, be misleading. Without considering changes across trials, for example, the analyses in Example 2 (Section 2.2) suggest that transparency of the second constituent has no effect and might have lead to the conclusion that semantic transparency is not involved in lexical processing or in written production (i.e., typing). These conclusions would be misleading because a model that takes into account changes across trials indicates that this variable does play a role.

Given these differences in interpretation for analyses with and without considering adaptation across trials, it would be useful to conduct analyses that evaluate this possibility, especially in situations in which a null effect would be interpreted as evidence that the linguistic construct (such as transparency) is not represented in the Mental Lexicon. Not all experiments will show adaptation, of course, and in these situations simpler models (i.e., models not involving by–trial interactions) are appropriate.

In terms of the literature on research methods and analysis, there have been many examples of how to deal with cross–trial changes, such as practice effects. Often this general speed–up is not relevant to the theoretical questions being addressed by the experiment (e.g., Heathcote, Popiel, & Newhort 1991) and, thus, a researcher might choose to remove such effects when analyzing the data or in the design of the experiment (Keppel & Zedeck, 1991). For example, counterbalancing of the experimental conditions can be used to equate the impact of practice effects across conditions. Note that the use of counterbalancing is not relevant to the experiments we presented because the items were randomized for each participants, which equates any general practice effects across conditions (as the conditions are distributed throughout the entire experiment). The selection of filler materials also influences practice effects (see Keuleers & Brysbaert 2011; Keuleers, Diependaele, & Brysbaert 2010).

There are several examples in the literature of temporal dependencies between successive trials (e.g., Broadbent 1971; Kinoshita, Mozer, & Forster
2011; Sanders 1998). For example, if the previous trial had a fast response then the subsequent trial might be slower. One method of dealing with trial–by–trial dependencies is to use time–series regression models in which autocorrelation functions are applied to the time series of response times (Cramer 1946; Hamilton 1997; Lutkepohl 2005; Wagenmakers, Farrel, & Ratcliff 2004; see also DeCarlo & Cross 1990 for an application of this technique to rating data). In essence, the response time on the preceding trial is used as a covariate. Doing so accounts for the variability associated with trial interdependences.

Although such techniques for including cross–trial effects are valid and are often used, they are not always appropriate because they might obscure relevant aspects of the data. These methods remove or minimize results that provide useful diagnostic information that could provide information that helps distinguish among various theoretical frameworks. Thus, before applying these methods, it is important to identify whether the changes across trials are attributable to non–theoretically relevant factors (such as general fatigue or practice with the task) or to theoretically relevant factors (such as fine–tuning of cognitive or language processing mechanisms). For example, sequential dependencies could be viewed as constraints on cognitive control and, consequently, trial–by–trial corrections are not something to be removed or controlled but rather are indicators of the way in which the operations of the cognitive system is fine–tuned during the experiment (see, for example, Mozer, Kinoshita, & Shettel 2007). Thus, systematic changes in responding across trials are not necessarily problems or irregularities in the data to be removed or corrected, but rather might be an important aspect of the data.

With respect to methods that correct for practice effects, it is important to note that these methods apply when the effects are constant across the various experimental conditions. For example, responses typically speed up across trials as the participant becomes more familiar with the task and the required response (e.g., knowledge about which buttons to press). If this increase is a true practice effect, it should not be dependent on the particular experimental condition. In our examples, we showed a condition by trial interaction which points to changes in the processing system and not to general practice effects.

Perhaps most critically, the results in our examples raise serious questions about cross–experiment comparisons and replications. One aspect of this problem that our examples raise is that the number of trials might influence the outcome of the experiment because the time that the system has to adjust differs. If the analysis only considers the main effect of a variable across trials, different experiments might yield different conclusions due to differences in the number of trials. This possibility is worth considering when comparing results across various experiments. In the psycholinguistic literature, there are several findings that seem to differ across experiments. To illustrate, Christianson, Johnson, and Rayner (2005) found that the naming time for compound words (e.g., sunshine) was faster relative to a control condition (e.g., sunsbine) when preceded by a prime for which the transposed letters occurred within a morpheme (e.g., sunhsine) but not when the transposed letters crossed the
morpheme boundary (e.g., susnhine). Duñabeitia, Perea, and Carreiras (2007) made a similar finding for lexical decision times for prefixed and suffixed words in Spanish and Basque; there was a transposed–letter effect when the transposition occurred within a morpheme, but not when the transposition crossed the morpheme boundary. In contrast, other researchers did find that transpositions across morpheme boundaries resulted in priming (e.g., Perea & Carreiras 2006; Rueckl & Rimzhim 2011). Another example of conflicting evidence in the literature concerns the influence of the frequency of constituents on the processing of compound words. Some studies have found that the frequency of both constituents influences ease of processing (e.g., Zwitserlood 1994). However, other studies have only found frequency effects of the first constituent (e.g., van Jaarsveld & Rattink 1988; Taft & Forster 1976), and still others have found effects only of the second constituent (Andrews 1986).

If the system is adapting over trials, and experiments differ in number of trials, the main effects in the experiments will include averages of drawn from trials involving different processing. That is, a longer experiment (i.e., one with more trials) will mean that the main effect will include more of the behaviour characteristic of the later stages of the adaptation, while a shorter experiment will only include the behaviour characteristic of the earlier stages of the adaptation. Hence, cross experiment differences may result simply due to the difference in the number of trials, even if exactly the same process of adaptation (and, the same processing in general) was occurring in the two experiments. This possibility is worth considering when researchers are trying to determine the cause of conflicting empirical results.

A second problem that might arise during cross–experiment comparisons is raised particularly by Example 4 (Section 2.4). Experimental replications very often include additional conditions (e.g., control conditions that arguably should have been included in a previous study) as a way of convincing reviewers that the replication is worthy of publication. Unfortunately, Example 4 indicates that in at least some circumstances the form that adaptation takes for a given condition can be strongly affected by the other conditions included in the experiment. It is particularly concerning, perhaps, that our inclusion of a completely unrelated baseline condition appears to have seriously affected the adaptation in the conditions of interest! In short, if the pattern of adaptation is sensitive to the other conditions that are included in an experiment, cross experiment comparisons must be made very carefully.

5. Concluding Remarks

In this paper, we have discussed the issue of adaptation in lexical processing and have provided examples of how failing to take into account adaptation during statistical analysis can lead to incorrect conclusions about the nature of the Mental Lexicon and about lexical processing in general. We argue that it is useful to examine potential interactions between the variable(s) of interest and trial because such interactions provide useful information about the
representations and processing that underly the responses. Moreover, we suggest that it is not always appropriate to use statistical techniques to remove changes across trials because not all changes are due to mere practice effects. Failing to consider the way in which trial interacts with the various experimental conditions can obscure insight into potentially interesting psychological processes. The scientific strategies that must adopted for studying an adaptive system are quite different from those used to study a static system. Although these strategies are not yet widely used or even well–developed in the context of psycholinguistic research, it is worth being aware of the phenomenon of adaptation so that researchers can seek out and develop new strategies to deal with this problem.

References


StataCorp (2013). *Stata Statistical Software: Release 13*. College Station: StataCorp LP.


Efekti prilagodbe u leksičkoj obradi

Psiholingvistička istraživanja obično primjenjuju znanstvenu strategiju koja pretpostavlja relativno stabilan skup predodžbi i procesa. U skladu s tom strategijom, istraživači mjere prosječne vrijednosti tako da izvode više pokusnih ispitivanja kako bi dobili statistički stabilnu procjenu uspješnosti izvedbe određenoga pokusnog uvjeta. U ovom radu izložit će se četiri skupa podataka izvadeni iz različitih psiholingvističkih zadataka i pokazati kako se psiholingvistički sustav naizgled prilagođava tijekom pokusnog ispitivanja. Pokazat će se kako postoje slučajevi u kojima određeni čimbenik nema glavni efekt, ali je u interakciji tijekom ispitivanja; u drugim slučajevima postoji glavni efekt čimbenika, ali je i taj čimbenik u interakciji s ispitivanjem. Konačno, pokazujemo slučajeve u kojima je način na koji čimbenik utječe na ispitivanje ovisan o drugim, uz njega nevezanim uvjetima pokusa. Naša se diskusija usredotočuje na teorijske i metodološke implikacije prilagodljivosti psiholingvističkog sustava.

**Key words:** compound words, linguistic adaptation, lexical processing, psycholinguistics

**Ključne riječi:** složenice, jezične adaptacije, jezično procesiranje, psiholingvistika