Est. 1841 | YORK ST JOHN UNIVERSITY

Hall, Christopher J ORCID logoORCID: https://orcid.org/0000-0001-9038-1238 (2018) The psycholinguistics of English. In: Hewings, A., Seargeant, P. and Pihlaja, A., (eds.) Routledge handbook of English Language Studies. Routledge

Downloaded from: https://ray.yorksj.ac.uk/id/eprint/2263/

The version presented here may differ from the published version or version of record. If you intend to cite from the work you are advised to consult the publisher's version: https://www.routledge.com/The-Routledge-Handbook-of-English-Language-Studies/Seargeant-Hewings-Pihlaja/p/book/9781138913455

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. Institutional Repository Policy Statement

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorksj.ac.uk

29 The psycholinguistics of English

Christopher J. Hall
York St John University

INTRODUCTION

Psycholinguistic studies have revealed a great deal about the nature of the English language, even though the primary objective of the field has been to explain the neuropsychological structures and mechanisms, which support the general human capacity for language. Psycholinguistics is concerned with three main issues. The first is the nature of mental representation of language: the organisation and storage of a person's language resources in memory, and how they are related to mentally represented conceptual knowledge, experience, and other cognitive capacities. The second is processing, concerning the mechanisms through which such resources are deployed in production and comprehension, through the modalities of speech, writing, and sign, and how these processes engage with individuals' non-linguistic knowledge, assumptions, and intentions. The third area, development, addresses the acquisition, ongoing modification (and possible attrition) of linguistic resources and processing mechanisms, as the result of innate capacities and social experience.

Psycholinguistics harnesses data from several sources to gain an understanding of these issues. In investigating processing, it has traditionally relied on behavioural data from controlled experiments, for example measuring the effects of prior exposure to semantically related words on target word recognition times ("priming") to investigate the organisation of the mental lexicon. To explore language development, experimental studies are complemented by observational data, where researchers have, for example, documented longitudinal changes in mean length of utterance in naturalistic settings. Increasingly, psycholinguistics is drawing on data from three sources which complement traditional methods. One source is the linguistic behaviour of people with language impairments, where psycholinguistics overlaps with speech-language pathology. Another source, related to impairment, is the measurement of brain activity, which has come to occupy a central role in recent years due to developments in neuro-imaging technology and other measures of neural responses to stimuli. A third source of evidence is computational modelling, particularly associated with connectionist theories of language development and use (see below).

Modern psycholinguistics has its roots in the "Chomskyan revolution" of the 1960s, when for the first time language was conceived theoretically as an innately determined cognitive capacity rather than as behavioural "habits" or as a structural system divorced from the mental life of individuals. In the early stages of psycholinguistic research there was an almost exclusive focus on the competence and performance of monolingual speakers of standard varieties of English. Indeed, in the words of Norris et al. (2001: 652), "[p]sycholinguistics has a long and embarrassing tradition of claims for language-universality based on data from English alone." Subsequently, in the latter decades of the last century, cross-linguistic work began to be conducted, using data from languages which are typologically distinct from English. Psycholinguistics has also emancipated itself from a strict adherence to innatist "formal" theories of language, exploring alternative functional models which revisit some behaviourist concepts, as subsequent sections illustrate.

The psycholinguistics of English can be viewed from several perspectives. One can, for example, focus on particular structures at different levels of linguistic organisation and examine how they are acquired, the nature of their mental representation, and the manner in which they are processed and selectively impaired. English has many properties in common with other languages, especially those with which it shares historical roots. However, at the same time it is unique in many respects, particularly in the way it has been moulded by sustained contact with other languages, and as a consequence has diverged significantly from the diachronic trajectories of its continental cousins. From this perspective, psycholinguistic data can shed light on the distinctive nature of English itself, as it is stored in, and deployed by, the minds of its users. Alternatively, one can separate the acquisition of English from the ways it is represented and processed once acquired, and the ways in

which it may be impaired. This is the way psycholinguistics as an academic field tends to be organised and pursued in scientific practice (i.e. in conferences and journals), with distinct (but increasingly overlapping) communities of scholars studying child development, second language acquisition (SLA), adult processing, and impairment. A third option is to approach the psycholinguistics of English from within particular theoretical orientations, and in what follows I indicate how English data have driven epistemological and ontological debate in the discipline.

Here, the focus will be on what research on mental representation, processing, and development reveal about English structure and usage, rather than, for instance, surveying milestones in the history of psycholinguistics which have been reached on the basis of English data. The focus is hard to maintain, however, given the primary goals of the discipline and its disproportionate reliance on English to test theory. Nevertheless, I have organised the chapter around psycholinguistic approaches to a selection of English structure and usage domains, giving a flavour of the methods used and how results have been used to support theory. The chapter begins with some observations about psycholinguistic perspectives on English, before exploring in more detail the development, representation and processing of a selection of English lexical and grammatical phenomena, in users and learners of different varieties of the language. Some current approaches to the ongoing debate about the psychological validity of rules are then outlined, followed by a brief discussion of future directions. The chapter ends with some general conclusions about the contribution of psycholinguistics to the field of English Studies.

CRITICAL ISSUES AND TOPICS

In this section I describe psycholinguistic work on a selection of English lexical and grammatical features which are distinctive to the language (e.g. the lack of transparency in the writing system), or have been subjected to intense scrutiny because of the light they shed on broader theoretical issues (e.g. the past tense *–ed* suffix). The section then moves on to address psycholinguistic processes in the Englishes of different populations of users, from speakers of regional dialects to adult second language learners.

Words

The development of vocabulary and the processing (production and recognition) of words have been extensively studied in English (cf., Aitchison, 2012). This section describes psycholinguistic work on (a) written and (b) morphologically complex words, drawing on data from first language development, word recognition and production, and language impairment. Studies of these two lexical-level domains of English have contributed significantly to theoretical debate in the discipline.

English is notorious among languages with alphabetic writing systems for its lack of consistent mapping between phonemes and letters, as reflected in educational debates about the use of phonics to teach children to read and spell conventionally, so processing the written word presents users of English with problems that are not as acute in other languages. A comparison of primary school children learning the orthography of 13 different European languages (Seymour et al. 2003) found that the participants learning English were over twice as slow as those acquiring more transparent systems like Finnish. Some psycholinguists have proposed that readers of English use two different routes to read a word: either by matching the whole word with its entry in the mental lexicon (the "direct route") or by using rules which match individual letters and letter groups (graphemes) with phonemes (the "indirect route") (Rastle 2007). For skilled readers, the direct route is the default and explains our ability to read opaque letter combinations (such as *island*) and those which are exceptions to rules (such as *have* compared with *cave, Dave, gave, nave, pave, rave, save, wave*). The indirect route is, of course, important for reading words we haven't encountered before and is therefore key to children's reading development. Proponents of "dual-route" models such as this propose that in normal reading, both routes are followed and factors like word frequency determine which route is faster.

Claims that the indirect route is psychologically real (i.e. not just a theoretical construct) are made in part based on evidence from the phenomenon of phonological dyslexia. People with this impairment cannot read letter sequences they do not know, often modelled experimentally using so-called *nonwords*: potential but non-occurring words in the language, like *tave*. They have no problem, however, reading words they know, whether they have regular or irregular/unpredictable spelling (e.g. *save*, *have*, *island*). According to the dual-route account, phonological dyslexics have an impairment of the indirect route, but the direct route is preserved. Conversely, people with a condition called surface dyslexia have trouble pronouncing written words that have irregular spelling (reading *island* as [ˈaislənd], for example) and are therefore argued to have an impaired direct route. Surface dyslexia is not attested in languages like Spanish or Serbian, which have a closer one-to-one mapping between graphemes and phonemes (cf., Harley, 2014: 226).

Other psycholinguists argue that the indirect route, although important for reading development, is unlikely to be involved in normal (unimpaired) reading. Moreover, they contend that other sound-letter mappings beyond grapheme-phoneme correspondences can be detected and used by readers. For example, children learning how to spell may, on repeated exposure to printed words, detect analogically that *–ave* is frequently pronounced [ev] across multiple words, and furthermore that the letter <a> in words like *save*, *hate*, and *came* consistently has the value /e/ when the letter <e> occurs after the postvocalic consonant. This has been simulated in computational psycholinguistics by connectionist models, which constitute complex networks of interconnected simple processing elements, like the neurons and synaptic connections of the brain. Such models, when "trained" with words of different frequencies, can "learn" how they are pronounced as a statistical learning problem, without any recourse to the kinds of grapheme-phoneme correspondence rules posited by the dual-route model (cf., Seidenberg, 2007). English sound and spelling thus provide a laboratory for broader debate about the psychological nature of human language.

Another structural domain of English words which has been studied in detail by psycholinguists and has fuelled intense debate about the psychology of language (and, indeed, about cognitive systems more generally) is inflectional morphology, particularly the past tense. Like the debate on how reading skills are developed and deployed, the central issue revolves around regularity, and again a dual-route mechanism is pitched against a single connectionist model (Ambridge and Lieven 2011). The default past tense form of an English verb is the stem plus the *–ed* suffix. However, there are around 180 verbs for which the past tense involves an alternation of the vowel (e.g. *throw*/*threw*), more substantial changes (e.g. *bring*/*brought*), or no change at all (e.g. *hit*/*hit*). Children develop their knowledge of English past tense forms following a U-shaped pattern. They first appear to learn both regular and irregular forms as wholes, one by one. Then, typically from their third year, they start to produce forms like *throwed*, even if they have correctly produced *threw* previously, suggesting they have constructed a rule (add *–ed*) which they tend to overgeneralise. Finally, from around six or so, their overgeneralisations cease and they settle on the adult system.

A second set of psycholinguistic data involving English past tense forms comes from people with language impairments (e.g. Ullman et al. 2005). As with dyslexia, there is evidence of a "double dissociation" between those who have trouble producing verbs with the regular past tense (e.g. in Broca's aphasia, an impairment of language production) and those with impaired access to irregular forms (e.g. in anomia, a word-finding deficit). Based on such evidence, some psycholinguists (notably Pinker, 1999), have argued for a dual-route model of English inflection, whereby direct access to inflected forms (e.g. *glided* and *threw*) exists alongside a rule system (add *–ed* for past tense). The rule system will be invoked in language production for regular verbs (*glided*) and newly-coined or newly-encountered verbs (e.g. *googled*), but not for irregular forms. For the production of very frequent regular forms, it is likely that the direct route will result in access to the whole form before the rule can assemble the stem and suffix. People with Broca's aphasia are assumed to have an impaired rule system, whereas those with anomia are assumed to have impaired direct access to irregular forms in the lexicon.

Proponents of connectionism, on the other hand, have argued that the regular/irregular dichotomy emerges from lower-level regularities in the data to which English speakers are exposed (cf. Joanisse and McClelland 2015). According to this position, children detect phonological and semantic patterns, often obtaining locally in a small numbers of verbs, with rule-like behaviour emerging only as an artefact of the statistical properties of the input. For example, the verb *throw* is much more likely to be over-regularised to *throwed* than *hit* is to *hitted*, because *throw* has several regularly-behaving phonological neighbours (e.g. *showed*, *flowed*, *rowed*, *sowed*) whereas *hit* has none. Neuropsychological evidence has been offered to counter the single-route model, showing that the processing of regular and irregular past tense forms is associated with distinct neuro-anatomical regions (cf. Marslen-Wilson 2007). The debate continues, however, with Westermann and Ruh (2012), for example, presenting a connectionist model which seeks to integrate findings from both the dual- and single-route paradigms.

Finally, English derivational morphology (the production of new words through, for example, affixation) presents similar mental challenges to those of spelling and inflectional morphology. There are large areas of phonological and semantic transparency, reflected in high productivity, coupled with considerable opacity. For processing, this yet again suggests the possibility of direct vs indirect routes in both production and comprehension: transparent morphologically complex words (e.g. *brave*/*bravely*) might be assembled online in production and parsed in comprehension, whereas opaque combinations (e.g. *broad/breadth*) might be produced and recognised as single units. Data from psycholinguistic experiments suggest that although derivational structure is detected and exploited during processing, the morphemes are not stored independently, except when transparent and productive (Marslen-Wilson 2007). Recently, neuropsychological data have been reported which suggest that English derivational forms are processed in the brain bilaterally (i.e. engaging both cerebral hemispheres), unlike regular inflection which depends on left hemisphere regions involved in broader grammatical processing (Bozic at al. 2013). There are, of course, competing connectionist accounts of English derivational morphology (e.g. Seidenberg and Gonnerman 2000), which stress the graded nature of this part of the language system.

Grammar

Compared with other languages, English has relatively fixed word order and little inflectional nominal morphology (only number). Processing sentences to find out "who did what to whom" depends to a large extent on global rather than local cues, for example, how the nouns are ordered with respect to the verb (argument structure), rather than case marking (e.g. nominal inflections for subject or object). In production, speakers and writers must formulate a sequential plan to indicate the thematic roles of the arguments of the main verb of a clause (agent, patient, etc.). Due to the absence of rich inflection in English, and the optionality of grammatical items such as complementisers like *that*, potential ambiguity is rife. One of the most famous English sentences in psycholinguistics illustrates this: *The horse raced past the barn fell* is very hard to process, even though it is grammatically well formed (cf. *The chocolate left in the sun melted*). English speakers expect the first noun to be the subject of the verb following it, whereas in these sentences it is the object, and the verb next to it is part of a passive relative clause (the main verb comes right at the end). In other languages, the role of *horse* and *barn* might be disambiguated with inflections, the past participle *raced* might have a different inflection from the simple past *raced*, and an overt complementiser might be required (*The horse <u>that was</u> raced past the barn fell*).

Children acquiring English are exposed early on to many tokens of subject-verb-object (SVO) constructions, in which the subject and object frequently correspond to agent-patient thematic roles (e.g. *Billy chased the dog*). This then becomes the "canonical" (default, most expected) order for the child. It has long been claimed that children are delayed in their grasp of the English passive construction, often until the sixth year, because it reverses the canonical order (e.g. *The dog was chased by Billy*). Psycholinguists differ, however, on whether the difficulty is uniquely grammatical. Maratsos et al. (1985) offered evidence that it is only 'non-actional' passives that are difficult for children (e.g. *Billy was <u>scared</u> by the dog* should be harder than *Billy was <u>bitten</u> by the dog*), and that

furthermore, short passives (without the *by*-phrase, e.g. *Billy was scared*) are easier than the full version. Two types of explanations have been offered, the first semantic, involving the degree of transitivity of the verbs involved, and the second syntactic, claiming that young children don't yet have access to complex grammatical principles for assigning thematic roles to noun phrases (NPs). But using syntactic priming, which involves measuring the extent to which individuals produce structures that they have just been exposed to, Messenger et al. (2012) showed that children as young as three do appear to have adult-like syntactic representations of the passive structure, independently of thematic role.

There are cases where English speakers prefer to produce passives rather than actives. The clearest cases are those where the object is animate. For example, speakers prefer to say *The man was hit by a truck* than *The truck hit the man*. This is in line with a general principle which favours earlier placement of animate over inanimate concepts in production (e.g. *I gave Kim a book* is preferred over *I gave a book to Kim*). Where both subject and object are animate, as in so-called reversible passives, processing is more taxing for adults, because the first noun in a sentence is more often an agent than a patient, and the sentence will make sense under this reading (i.e. *The girl was pushed by the boy* makes as much sense as *The girl pushed the boy*). People with impaired grammatical ability, such as those with Broca's aphasia, interpret reversible passives correctly at only chance level, compared with much higher levels of accuracy with non-reversibles. This may be due to their insensitivity to grammatical items with little inherent semantic content, such as English passive *be* V-*ed* (e.g. Ingram 2007: 251-253).

Another group that perform poorly with reversible passives are children with Specific Language Impairment (SLI) (Leonard 2014: 78). This neurological disorder affects language but is not attributable to brain damage, hearing loss, or intellectual disability. Since it appears to run in families, there has been much speculation (and many news stories) about a possible genetic cause. One English-speaking family of three generations, in which around half the family members exhibited symptoms of SLI, has been studied in depth, with initial reports stressing the specificity of the impairment and later ones questioning it (Gopnik, 1990; Vargha-Khadem et al. 1995). A gene called FOXP2 (the so-called "grammar gene") emerged as a major factor in the development and transmission of the disorder. People with SLI do have disordered grammar, including past tense *–ed*, which they appear able to formulate a rule for, yet apply very inconsistently.

Unstandardised Englishes

It is a regrettable fact that the stimuli used in psycholinguistic experiments and language tests almost always assume the norms of "Standard English", with the presence of unstandardised varieties generally ignored. An area of applied psycholinguistics that does focus on such varieties is literacy development, given that for speakers of regional and social accents, phoneme-grapheme correspondences are even less transparent than for users of standardised varieties (from which written norms originally developed). Many educationalists assume that this contributes to poorer literacy skills among speakers of unstandardised dialects. Yet recent psycholinguistic research (e.g., Terry 2014) presents evidence from experimental tasks suggesting that for African American English (AAE) speakers, it is their degree of phonological awareness, rather than the simple presence of dialect features, which accounts for their performance on tests of standard American English. Related to this, speech language therapists have often pointed out that speakers of unstandardised dialects are at risk of being diagnosed with language impairments, given their enhanced likelihood of scoring poorly on tests of metalinguistic knowledge of standardised forms. One study compared white and African American children who were either language-impaired or typically developing, using tests of metalinguistic knowledge vs. tests of processing ability (Rodekohr and Haynes 2001). They found that both kinds of test distinguished impaired from typically-developing children, but that on the knowledge test the AAE speakers performed significantly closer to the impaired children. They conclude that the use of processing tests can reduce bias against unstandardised dialect users in language screening.

Many monolingual children grow up with multidialectal input, for example when each parent speaks a different variety of English. Some psycholinguists have argued that such children develop in ways similar to early bilinguals, who have been demonstrated to construct less specific initial mental representations for word forms than their monolingual peers do. Durrant et al. (2015) tested this using a task in which infants of around 20 months heard words which matched one of a pair of images (e.g., they hear *cup* and see images of a cup and a clock), and their gaze direction and fixation were measured. In some trials the target was mispronounced (e.g. [gʌp] or [kɛp] instead of [kʌp]). The monodialectal children (both of whose parents spoke British Southwest English dialect) did not gaze at the target image after a mispronunciation, whereas children with parents who spoke two different dialects *did*. This suggests that children exposed to two dialects have less specific representations for word forms in their mental lexicons, such that divergences in pronunciation can be accommodated.

How (more commonly, how well) individuals can acquire a second dialect after their first is another topic which has seen expanding interest in recent years. Like SLA, work in this area examines the processes involved in naturalistic and educational contexts (the latter concentrating on the development of standardised varieties), but tends to focus on accent at the expense of other linguistic features. Siegel (2010) provides a survey of eight studies of naturalistic English SDA, in which most of the features studied were phonological. There are very few cases in which individuals reached above 90% of target feature usage, and these are amongst the youngest learners studied. Finally, there are several studies of speech intelligibility within and across dialects, which often appear to show that some dialects are inherently more intelligible than others. McCloy et al. (2015), however, using more stringent controls than previous studies, conducted a perception task in which speakers of two different northern US dialects listened to tokens of each dialect in different levels of background noise. The results reveal the difficulty of teasing apart the effects of dialect features from idiolectal (individual) variation — a factor we return to subsequently.

English as a second language

There are now many more learners and non-native users of English than native speakers (NSs). Second language (L2) Englishes and how they develop is therefore a major area of interest for psycholinguists. The modern discipline of SLA can be traced to a dramatic shift in interest from practical issues of teaching to cognitive theories of learning, following the Chomskyan revolution. Although much research in the discipline now has a social, and indeed often an overtly *anti*-cognitive, orientation, cognitive issues are still central to SLA theory. As in other areas of psycholinguistics, English has been the dominant language studied, for the most part as a goal of adult learning in instructional contexts. But because English speakers can be L2 learners too, researchers have also collected substantial amounts of data on the development of other languages (notably Spanish in the USA and French in the UK and Canada). Indeed, two of the most psycholinguistically-oriented theories in SLA are the result of research on languages other than English. Processability Theory (cf. Pienemann and Lenzing 2015), originally developed on the basis of studies of learners of German, claims that the course of language development is constrained by what the learner can process at any given stage in their learning trajectory. For example, learners of English will only be ready to acquire the subject-auxiliary inversion construction for questions after they have learnt that questions are formed with an initial *wh*- word. Before then, they will produce utterances like *Where he has been?* Similarly, Input Processing (IP) theory (cf. VanPatten 2015), based originally on data from English-speaking learners of Spanish, aims to explain how principles of form-meaning mapping in comprehension can influence the emerging L2 system. One IP principle states that grammatical forms with more information content will be processed earlier than more redundant forms, e.g. English *–ing*, which marks progressive, over third person *–s* (one of the hardest forms for learners of L2 English to master).

A major focus of study on L2 Englishes in SLA and psycholinguistics more generally has been the ways in which the first language of learners influences the second. Studies of transfer, or more precisely cross-linguistic influence (CLI), have been conducted at all levels of language knowledge, from pronunciation to pragmatics. At the pronunciation level, “foreign accent” is the norm for L2

English users, characterised by the use of phonemes like /d/ and /z/ for /ð/ and their unvoiced versions (e.g. *dese tings* or *zese sings* for *these things*). However, English is no different from other languages in being susceptible to CLI at this level, given that phonology is the earliest established component of linguistic knowledge in infancy and is particularly hard to displace. At the lexical level, CLI will vary according to the degree of overlap between learners' L1 and English. Adult learners whose L1 has inherited word forms from a common Germanic ancestor (such as Dutch) or has contributed word forms to English following conquest and cultural influence (such as Latin and its descendants), will find English vocabulary easier than L1 speakers of unrelated languages like Mandarin Chinese or Arabic. There is a great deal of evidence to suggest that cognate forms are automatically activated in the processing of L2 words, often leading to an assumption of translation equivalence (cf. Hall, 2002). This happens even when it is not warranted, as in the case of so-called "false friends" like Spanish *tuna* ('prickly pear') being taken as the translation equivalent of English *tuna*.

Research on the mental lexicons of bilingual speakers and learners of English has convincingly shown that the two vocabularies are stored in a single interconnected network, even at high levels of proficiency. A similar consensus has emerged more recently about grammatical resources and processing, and evidence is accumulating for co-activation of both languages in performance irrespective of the typological relatedness of the other language or the direction of CLI (L1→L2 or L2→L1). In a review article, Kroll et al. (2012) refer to evidence that bilinguals whose other language is completely unrelated to English (even American Sign Language) still activate it when reading English, and that bilingual Spanish NSs who live in English-dominant environments are influenced by English relative clause processing strategies when reading Spanish. Some grammatical constructions in English appear to cause problems for L2 learners because they do not exist in the L1, so no CLI is possible, either positive or negative. An example is the article system, typified by the distinction between indefinite *a* and definite *the*, used to determine NP reference. Chinese lacks articles, and Chinese-speaking learners of English often use them inconsistently in production. But a study using eye-tracking (Trenkic et al. 2014) reveals that in comprehension, intermediate proficiency learners, like NSs, used the information in articles as well as contextual information to disambiguate NP reference (albeit not as fast as the NSs). This evidence converges with the results of an increasing number of studies which suggest that L2 learners/users can process grammatical structures essentially the same way as NSs, using the same areas of the brain (e.g. Abutalebi and Della Rosa, 2012).

KEY AREAS OF DISPUTE AND DEBATE

We have inevitably already touched on the central debate in psycholinguistics, namely whether the Chomskyan notion of rules operating on symbols is psychologically valid, or whether connectionism provides a more plausible explanatory framework. In this section, we explore three different ways in which this debate has been taken forward.

Usage-based approaches

Frequency is a key variable in psycholinguistics at the lexical level. There is robust evidence, for example, that infrequent words like *milt* (occurring twice per million) take longer to recognise than more frequent but otherwise similar words like *milk* (occurring 49 times per million). Yet up until relatively recently psycholinguists have not wanted to assign frequency a broader role in the psychology of language, because it evokes the idea of 'habit formation' associated with the behaviourist paradigm discredited by Chomsky. According to the Chomskyan position, language is an innate cognitive faculty constrained by a Universal Grammar (UG) of rules governing well-formed strings of symbols. It is UG that allows children to extract patterns from the input, which is seen to be inherently inadequate on its own. An alternative view in which input frequency is key has been gaining strength over the past decade or so in a family of theoretical approaches called Usage-based Linguistics (UBL). According to UBL, language is a cognitive resource constructed and continuously developing on the basis of individual users' analyses of the frequency and distribution of form-meaning pairings in the input experienced during usage events (Tomasello 2003). Instead of UG, learners use general-purpose procedures like analogy. Take the English passive, for example.

Ambridge and Lieven (2011: 279) observe that "children begin with lexically specific […] frames (e.g. *it's broken*) that become increasingly abstract (e.g. → *it got VERBed by it* → *NP BE/GET VERB by NP*)". The approach has also been extended to processing. MacDonald (2013) provides a particularly compelling usage-based framework for explaining processing phenomena as well as numerous facts about English grammar by showing how processing demands in production have helped determine the frequency and distribution of constructions like the passive and relative clauses, which in turn mould comprehension strategies.

The UBL approach has been applied to L2 Englishes also. Wolter and Gyllstad (2013), for example, show that Swedish learners of English are sensitive to the frequencies of collocations in English. But they also point out that CLI still plays a major role. This is confirmed also by Hall et al. (2017), who investigated the ways in which input interacts with other learning factors to shape lexico-grammatical features in the idiolect of a single expert user of L2 English. The UBL approach is likely to have particularly dramatic effects on SLA theory and our understanding of L2 Englishes in coming years, perhaps overturning several long-held assumptions. One emerging possibility is that many of the purported differences between native and non-native speakers are essentially the same as those observed *between natives*. Kaan (2014), for example, highlights the role of input frequency, the amount of competing forms, and the variable quality/consistency of lexical knowledge, as common factors which predict the degree to which both native and non-native users make predictions of upcoming material in online comprehension. Her position is completely consistent with UBL approaches to language development.

Individual differences

Although there are several (and some very well-known) case studies of first and second language development and impairment (e.g. Brown, 1973), most studies in psycholinguistics are based on behavioural or neural activity data measured across groups, with results averaged out and taken as characterisations of whole populations (such as English NSs). Consistent with this approach, computational models simulate "prototypical" development and processing. However, a growing interest in differences between individuals is now evident, often related to social experience, especially educational level. Street and Dąbrowska (2014), for example, report data which they interpret as showing that English-speaking individuals with lower academic attainment have greater problems in comprehending the passive that those with higher education. They explain this by appealing to UBL principles: "Passive sentences are considerably more frequent in written texts […] than in speech; because more educated participants tend to read more, their passive constructions are better entrenched, and hence accessed more reliably, which results in faster and more accurate performance" (p. 113). Similarly, a study by Yap et al. (2012) using word recognition data from over 1,200 participants in the English Lexicon Project (see below), revealed marked stability within individual performance but substantial variability between individuals, correlated significantly with vocabulary size. Finally, Paradis (2011) examined the role of individual variables in a study of over 160 child acquirers of L2 English. He found that internal factors such as short-term memory capacity were better predictors of vocabulary size and accuracy with verbal morphology than external factors such as the richness of the input environment. Like Hall et al. (2017), he concludes that his findings are consistent with UBL, but that input is not the whole story. All the evidence so far suggests that, contrary to the UG position, the development of English (and by extension other languages) is not fully constrained by universal principles.

The phonics debate

Psycholinguists have not normally used their findings to influence applied linguistic practice, although their research has significant impact on applied issues like speech language pathology and TESOL. One arena in which they have advocated specific public policy measures is in the teaching of reading. Traditionally, the learning of letter names and their associated sounds was emphasised in teaching ("A is for *apple*, B is for *book*"; "D-O-G spells *dog*"; etc.). In the 1960s there was a shift to "whole-word" teaching which de-emphasised the alphabetic principle, justified largely on the basis of

the irregularity of English spelling, and later "whole-language" teaching, which emphasised the top-down processing of text meaning over the bottom-up processing of word forms. Educationalists referred extensively to Goodman's (1967) suggestion that reading was a "psycholinguistic guessing game," in which strategies for extracting meaning (from context, including pictures) were as important as rule-based processing of letter combinations. A couple of decades later, many psycholinguists (especially in the USA) started to voice their concern with "whole-word" and "whole-language" methods, vigorously supporting calls for a return to the teaching of grapheme-phoneme correspondence rules in the approach known as "phonics". According to Rayner et al. (2001: 56): "The main rationale behind a phonics approach is that it explicitly teaches children both the alphabetic principle and the specific letter-phoneme correspondences that generalize across many English words." They claim that phonics is consistent with both dual-route and connectionist models. But the pendulum continues to swing, with numerous psycholinguists now questioning phonics instruction and the assessment regimes that come with it. In the UK context, for instance, Gibson and England (2015) argue that the use of non-words in statutory phonics testing for primary school pupils is problematic. They express concerns about: (a) the validity of the test items themselves and teachers' ability to discriminate all plausible responses to them; (b) the assumption that non-word reading performance involves the same mental processes as normal reading; and (c) the evidence that orthographically opaque languages like English lead to significant delays in reading development in the first place.

FUTURE DIRECTIONS

Psycholinguistic studies of English will continue to benefit from new developments in neuropsychology, which should allow more fine-grained analyses of cognitive representation and processing for language. Two areas of future growth in the discipline which are rather more specific to English are: (a) the use of large computerised databases of language in authentic usage and of norms for specific words or constructions from psycholinguistic studies (e.g. average word recognition times); and (b) the recognition that English is now most typically processed and mentally represented as part of a multilingual repertoire, often by NNSs.

So-called "megastudies" (cf. Keuleers and Balota 2015), originally developed for research on visual word recognition, use databases recording behavioural measures on several psycholinguistic variables, such as word frequency or spelling consistency, drawn from very large collections of words and individuals. They complement the traditional experimental approach that typically involves the comparison of participant behaviour on a small set of stimuli representing usually two variables to be correlated, selected by the experimenter. Megastudy data have been used to investigate several variables, including those just mentioned, as well as age of acquisition and imageability (how easily a word is associated with a visual image). The biggest megastudy database is the English Lexicon Project (freely available at http://elexicon.wustl.edu) which includes millions of measurements on over 40,000 words from several hundred participants. Crowdsourcing is also being increasingly used to obtain psycholinguistically relevant data on English. We have seen that word frequency is an important variable in processing, however another significant measure is *word prevalence*, i.e. how many English speakers know the word, which can be measured through crowd-based experiments.

Work in applied linguistics on non-native speaker Englishes has recently begun to have an impact on mainstream linguistics and the ways in which linguists have tended to conceive of named languages as "monolithic" entities, identified with monolingual speakers of standardised varieties. Although there is a great deal of psycholinguistic work on bi- and multilingual users of English, much of it deals with child or adult learners, rather than non-native *users*. With increased global movement of people and emphasis on English competence in national education policies, this group of English users is set to grow massively. Increasingly, their usage of English will be in lingua franca contexts, i.e. as a bridge language with people who have a different first language. In such contexts, shared group norms will inevitably be attenuated, so the manner in which they process English, and the ways in which their mental lexicons and grammars develop as a result of such post-instruction usage, are of considerable psycholinguistic interest (cf. Hall forthcoming).

CONCLUSION

Psycholinguistics has been informed by English data from the earliest days of its establishment as a discipline, with features of the language providing test cases for the development of theoretical models of language cognition. Yet psycholinguistics has also informed English Studies, especially in recent decades, by revealing how the language presents cognitive challenges which differ from those faced by users of other languages. Current approaches to psycholinguistics are less exclusively reliant on the precepts of theoretical linguistics than they were in the heyday of the Chomskyan revolution, with the result that the discipline has even greater potential to contribute to broader interdisciplinary enterprises like English Studies. When combined with corpus linguistics, for example, psycholinguistics can contribute to our understanding of variation and change in the language (e.g. Gries 2013). A psycholinguistic perspective therefore complements other approaches to English by revealing how, in addition to being a linguistic system, a community resource, and a potent marker of identity, it is also a cognitive resource which is developed, stored, and processed in individual minds.

FURTHER READING

Bialystok, E., Craik, F. I., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240-250. (This is a comprehensive review of issues in the psycholinguistics of bilingualism, using many examples from bilingual users of English and stressing the cognitive advantages of being bilingual.)

Harley, T. (2014). *The psychology of language* (4th edn.). London: Psychology Press. (This textbook is an impressive compendium of empirical findings and theoretical models in psycholinguistics, drawing mostly on English data.)

Ingram, J. C. (2007). *Neurolinguistics: An introduction to spoken language processing and its disorders*. Cambridge: Cambridge University Press. (A more narrowly neurolinguistic view of processing and impairment, also based mostly on English data.)

Sparks, J. R., and Rapp, D. N. (2010). Discourse processing—examining our everyday language experiences. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(3), 371-381. (An overview of issues in discourse processing, again using mostly English examples.)

Vitevitch, M.S., Chan, K.Y. and Goldstein, R. (2014). Using English as a 'model language' to understand language processing. In Miller, N. and Lowit, A. (eds.), *Motor speech disorders. A cross-language perspective* (58-73). Bristol: Multilingual Matters. (An overview of phonological processing phenomena which are distinctive in English, in both normal and impaired language users.)

RELATED TOPICS

6, 7, 8, 12, 30

REFERENCES

Abutalebi, J. and Della Rosa, P. A. (2012). How the brain acquires, processes, and controls a second language. In Faust, M. (ed.), *The handbook of the neuropsychology of language* (pp. 516-538). Oxford: Blackwell.

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon* (4th edn). Chichester: Wiley-Blackwell.

Ambridge, B. and Lieven, E.V. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge: Cambridge University Press.

Bozic, M., Tyler, L.K., Su, L., Wingfield, C. and Marslen-Wilson, W.D. (2013). Neurobiological systems for lexical representation and analysis in English. *Journal of Cognitive Neuroscience*, *25*(10), 1678-1691.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
Durrant, S., Delle Luche, C., Cattani, A., and Floccia, C. (2015). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language, 42*(02), 447-465.
Gibson, H. and England, J. (2015). The inclusion of pseudowords within the year one phonics 'Screening Check'in English primary schools. *Cambridge Journal of Education*, 1-17.
Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Literacy Research and Instruction, 6*(4), 126-135.
Gopnik, M. (1990). Feature-blind grammar and dysphasia. *Nature, 344*(6268), 715.
Gries, S. T. (2013). Sources of variability relevant to the cognitive sociolinguist, and corpus- as well as psycholinguistic methods and notions to handle them. *Journal of Pragmatics, 52*, 5-16.
Hall, C. J. (2002). The automatic cognate form assumption: Evidence for the Parasitic Model of vocabulary development. *International Review of Applied Linguistics, 40*, 69-87.
Hall, C. J. (forthcoming). Cognitive perspectives on English as a Lingua Franca. In Jenkins, J., Baker, W. and Dewey, M. (eds). *Routledge handbook of English as a Lingua Franca*. London: Routledge.
Hall, C. J., Joyce, J. and Robson, C. (2017). Investigating the lexico-grammatical resources of a non-native user of English: the case of *can* and *could* in email requests. *Applied Linguistics Review, 8*,(1), 35-59.
Harley, T. (2014). *The psychology of language* (4th edn.). London: Psychology Press.
Ingram, J. C. (2007). *Neurolinguistics: An introduction to spoken language processing and its disorders*. Cambridge: Cambridge University Press.
Joanisse, M. F. and McClelland, J. L. (2015). Connectionist perspectives on language learning, representation and processing. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(3), 235-247.
Kaan, E. (2014). Predictive sentence processing in L2 and L1. What is different? *Linguistic Approaches to Bilingualism, 4*, 2, 257-282.
Kroll, J. F., Dussias, P. E., Bogulski, C. A., and Valdes Kroff, J. R. (2012). Juggling two languages in one mind: what bilinguals tell us about language processing and its consequences for cognition. *Psychology of Learning and Motivation, 56*, 229-262.
Keuleers, E., and Balota, D. A. (2015). Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments. *The Quarterly Journal of Experimental Psychology, 68*(8), 1457-1468.
Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT press.
MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4*, 226, 1-16.
Maratsos, M., Fox, D., Becker, J., and Chalkley, M. (1985). Semantic restrictions on children's passives. *Cognition, 19*, 167–191.
Marslen-Wilson, W.D. (2007). Morphological processes in language comprehension. In Gaskell, G. (ed.), *The Oxford Handbook of Psycholinguistics*, pp. 175-193. Oxford: Oxford University Press.
McCloy, D. R., Wright, R. A., and Souza, P. E. (2015). Talker versus dialect effects on speech intelligibility: A symmetrical study. *Language and Speech, 58* (3), 371-386.
Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language, 66*(4), 568-587.
Norris, D., McQueen, J. M., Cutler, A., Butterfield, S. and Kearns, R. (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes, 16* (5/6), 637-660.
Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*(3), 213-237.
Pienemann, M. and Lenzing, A. (2015). Processability theory. In VanPatten, B. and Williams, J. (eds.), *Theories in Second Language Acquisition: an introduction* (2nd edn; pp. 159-179). London: Routledge.
Pinker, S. (1999). *Words and rules. The ingredients of language.* New York: Basic Books.
Rastle, K. (2007). Visual word recognition. In Gaskell, G. (ed.), *The Oxford Handbook of Psycholinguistics*, pp. 71-87. Oxford: Oxford University Press.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., and Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*(2), 31-74.

Rodekohr, R. K. and Haynes, W. O. (2001). Differentiating dialect from disorder: A comparison of two processing tasks and a standardized language test. *Journal of Communication Disorders*, *34*(3), 255-272.

Seidenberg, M.S. (2007). Connectionist models of reading. In Gaskell, G. (ed.), *The Oxford handbook of psycholinguistics*, pp. 235-250. Oxford: Oxford University Press.

Seidenberg, M.S. and Gonnerman, L.M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, *4*(9), 353-361.

Seymour, P. H., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143-174.

Siegel, J. (2010). *Second dialect acquisition*. Cambridge: Cambridge University Press.

Street, J. A. and Dąbrowska, E. (2014). Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive. *Applied Psycholinguistics*, *35*(01), 97-118.

Terry, N. P. (2014). Dialect variation and phonological knowledge: Phonological representations and metalinguistic awareness among beginning readers who speak nonmainstream American English. *Applied Psycholinguistics*, *35*, 155-176.

Tomasello, M. (2003). *Constructing a language. A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Trenkic, D., Mirkovic, J., and Altmann, G. (2014). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition*, *17*(2), 237-257.

Ullman, M.T., Pancheva, R., Love, T., Yee, E., Swinney, D. and Hickok, G. (2005). Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language*, *93*(2), 185-238.

VanPatten, B, (2015). Input processing in adult SLA. In VanPatten, B. and Williams, J. (eds.), *Theories in Second Language Acquisition: an introduction* (2nd edn; pp. 113-134). London: Routledge.

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P., & Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences*, *92*(3), 930-933.

Westermann, G. and Ruh, N. (2012). A neuroconstructivist model of past tense development and processing. *Psychological Review*, *119*(3), 649-667.

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition*, *35*(03), 451-482.

Yap, M. J., Balota, D. A., Sibley, D. E., and Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 53.

Word count: 7635

Biographical note

Chris Hall is Professor of Applied Linguistics in the School of Languages and Linguistics at York St John University. His publications include *Morphology and Mind*, *An Introduction to Language and Linguistics*, *Mapping Applied Linguistics*, and *Introducing Language in Use*.