Spoken word recognition in adolescents with autism spectrum disorders and specific language impairment

TOM LOUCAS and NICK RICHES
University of Reading

GILLIAN BAIRD
Guy’s Hospital, London

ANDREW PICKLES and EMILY SIMONOFF
De Crespigny Park, London

SUSIE CHANDLER and TONY CHARMAN
Institute of Education, London

Received: March 12, 2010    Accepted for publication: April 21, 2011

ADDRESS FOR CORRESPONDENCE
Tom Loucas, School of Psychology and Clinical Language Sciences, University of Reading, Earley Gate, Reading RG6 6AL, United Kingdom. E-mail: t.loucas@reading.ac.uk

ABSTRACT
Spoken word recognition, during gating, appears intact in specific language impairment (SLI). This study used gating to investigate the process in adolescents with autism spectrum disorders plus language impairment (ALI). Adolescents with ALI, SLI, and typical language development (TLD), matched on nonverbal IQ listened to gated words that varied in frequency (low/high) and number of phonological onset neighbors (low/high density). Adolescents with ALI required more speech input to initially identify low-frequency words with low competitor density than those with SLI and those with TLD, who did not differ. These differences may be due to less well specified word form representations in ALI.

The first stage in constructing an utterance interpretation involves mapping a representation of acoustic–phonetic input onto representations of word form stored in the mental lexicon. This process of lexical access provides the logical starting point for an investigation of how listeners process spoken language. Probing spoken word recognition in children with language impairments may shed light on the nature of the deficits in the language processing system that characterize many children with developmental disorders.

© Cambridge University Press 2011 0142-7164/11 $15.00
Specific language impairment (SLI) and autism spectrum disorders (ASDs) are common developmental difficulties associated with language impairment. SLI is characterized by deficits in the structural aspects of language that cannot be explained by hearing impairment, intellectual disability, frank neurological impairment, or ASD. In contrast, ASD is a pervasive developmental disorder that is characterized by impairments in reciprocal social interaction and communication and by restricted and repetitive behaviors and interests. Pragmatic impairments have driven research into language and communication in ASD, as difficulties in this area are almost universal and found regardless of level of intellectual functioning, but structural language impairments are also associated with ASD (for a review, see Tager-Flusberg, Paul, & Lord, 2005). A subgroup of children with ASD presents with structural language impairments in the context of nonverbal skills within the average range (henceforth, ALI). This psychometric profile typical of SLI invites the comparison between SLI and ALI (Tager-Flusberg, 2006).

Deficits can present at all structural levels of language in SLI, including the lexicon. First words are delayed in children with SLI, estimated to emerge at 23 months compared with 11 months in typically developing children in Trauner, Wulfek, Tallal, and Hesselink’s (2000) sample. Children with SLI continue to lag behind age-matched peers on receptive vocabulary measures (Rice, 2004). Many children with SLI also have word-finding difficulties (Dockrell, Messer, George, & Wilson, 1998) and perform poorly on word-definition tasks (Dockrell, Messer, George, & Ralli, 2003). It has consequently been suggested that semantic knowledge associated with words may be underspecified in these children’s lexicons (McGregor, Newman, Reilly, & Capone, 2002). Similarly, in ASD, deficits at the lexical level are apparent. In children with ALI first words emerge late, at 22 months in Loucas et al.’s (2008) population-based sample. However, unlike children with SLI, a significant minority of children with ASD, some 15%, acquire first words without any apparent delay and subsequently lose language skills in the second year of life (Pickles et al., 2009). The literature on word-finding difficulties in ASD is limited, although Rapin and Dunn (2003) report that higher order language impairments, which include word-finding difficulties, are more common in children with ASD than those with SLI. Unlike those with SLI, children with ASD may also use words idiosyncratically, an observation made by Kanner (1946) in the first report of autism. This use of ordinary words with special or unique meanings, not shared by others and the coining of new words (neologisms), is found in individuals with ASD regardless of developmental level, and suggests that it does not mark a stage in lexical acquisition (Volden & Lord, 1991). Thus, when comparing SLI and ASD, there are both similarities and differences in broad aspects of the lexical knowledge and processing. The study presented here aims to probe in detail the early stages of spoken word recognition to investigate possible similarities and differences in underlying aspects of lexical representation and process in SLI and ALI.

Adult listeners are able to recognize spoken words rapidly and efficiently. In context, recognition can occur before sufficient acoustic phonetic information has become available to allow correct identification on that basis alone; that is, with around 250 ms of input, or typically only two phonemes into a word, when on average 40 words would match these two phonemes (Marslen-Wilson, 1987). This
is achieved by partial activation of all of those lexical representations that match input as it unfolds over time, which compete for recognition until the choice is made as the candidate emerges from the activation levels of competitors (Marslen-Wilson, 1990). Current models of spoken word recognition differ in how lexical competition is defined. Some, such as the Cohort model (Gaskell & Marslen-Wilson, 1997; Marslen-Wilson & Welsh, 1978) emphasize word onsets, while others, such as TRACE (McClelland & Elman, 1986) and neighborhood activation model (Luce & Pisoni, 1998), emphasize global similarity. The properties of both word onsets and neighborhoods effect the dynamics of spoken word recognition. Although the number, or density, of onset competitors has a continuous inhibitory effect, neighborhood density appears to have early facilitatory effects and late inhibitory effects (Magnuson, Dixon, Tanenhaus, & Aslin, 2007).

From a developmental perspective there are a number of issues that arise when considering how the properties of the adult system emerge during the course of development: whether lexical competition characterizes spoken word recognition in children, whether the lexicon is organized around the phonological form of lexical representations such that competitors can be understood in terms of the phonological similarity to targets, whether word-form representations show a level of representational detail that allows a high degree of fit between acoustic phonetic input and stored representations. Evidence from studies of spoken word recognition in children suggests that similar processes are indeed at work during the development of the process.

The properties of spoken word recognition can be tracked using the gating task in which listeners are presented with increasingly longer fragments of a word (gates) and required to identify the word after each gate. Gating studies suggest that children can recognize words on the basis of partial input, but they need to hear more acoustic–phonetic information to identify items than adults (Elliot, Hammer, & Evan, 1987). Metsala (1997) investigated the effects of word frequency and neighborhood density, finding that adults and 11-year-old children required less acoustic–phonetic information to recognize low-frequency words and words in sparse neighborhoods than 7- and 9-year-olds. Metsala (1998) concluded that this development pattern could be explained if children’s word form representations are less well specified than those of adults, arguing that word form representations only become well specified gradually under the pressure of increasing vocabulary size and the need to differentiate between minimally different forms. The rate at which a word-form representation is restructured depends on its frequency of occurrence and phonological distinctness from its neighbors. Thus, developmental differences were most pronounced for low-frequency words with few neighbors because they are under least pressure to restructure.

The development of spoken word recognition has also been investigated in children with atypical language development. Studies using the gating paradigm, have found that children with SLI can recognize spoken words from the same amount of acoustic–phonetic information as typically developing children matched for vocabulary (Dollaghan, 1998; Montgomery, 1999), although children with SLI may require more input to recognize newly taught words (Dollaghan, 1998). Mainela-Arnold, Evans, and Coady (2008) investigated the hypothesis that children with SLI have less precise word form representations than typically developing children.
using Metsala’s (1998) manipulation of frequency and neighborhood density. They predicted if children with SLI have less well-specified lexical representations they would show a smaller effect of neighborhood density and less efficient recognition of low-frequency words. However, children with SLI recognized words with similar amounts of input as age-matched controls. They nonetheless found that for late gates children with SLI were less certain about their lexical choices than controls, which they speculate may be due to weak lexical representations that are not captured in the frequency and neighborhood manipulations used in their study.

Much of the gating evidence suggests that spoken word recognition is relatively unimpaired in children with SLI. Mainela-Arnold et al. (2008) argued that, as children with SLI are input dependent, for example, requiring more presentations to acquire a new word (Rice, Oetting, Marquis, Bode, & Pae, 1994), their performance on lexical access tasks such as the gating task would demonstrate a reduced effect of frequency and competitor density. Insofar as this hypothesis was not supported by their data, this may reflect relatively intact input processing, and indeed, the evidence for impaired speech processing in SLI is at best equivocal (Ellis Weismer, 2005). Spoken word recognition has not been investigated in ASD. However, there is evidence of abnormal speech processing in ASD that may lead to greater input dependence as hypothesized in SLI, which may emerge as deficits in spoken word recognition. Children with ASD show impaired involuntary orienting to speech (Kuhl, Coffey-Corina, Paden, & Dawson, 2005) and abnormal responses to changes in the speech signal (Ceponiene et al., 2003). In ASD, these speech-processing difficulties may lead to an input dependence that attenuates the formation of well-specified representations. This would be most evident where the constraints driving the development of word form representations are weakest, that is, words that are low in frequency and have few competitors. Thus, the present study asks if the processes of spoken word recognition differ in adolescents with ALI compared with adolescents with SLI. In addition, we seek to replicate Mainela-Arnold et al.’s (2008) findings for SLI using a different definition of competitor density: onsets rather phonological neighborhoods.

METHODS

Participants

The study investigated two clinical populations: adolescents with SLI and adolescents with ALI. Twenty-six adolescents with SLI or ALI were recruited from a cohort of children with Special Educational Needs who had been assessed during the Special Needs and Autism Project (SNAP; Baird et al., 2006). A diagnosis of autism was made on the basis of ICD-10 (World Health Organization, 1993) criteria using the Autism Diagnostic Interview—Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994), the Autism Diagnostic Observation Schedule—Generic (ADOS-G; Lord et al., 2000), and additional information from local clinical assessment and from schools (full details of the diagnostic process are available in Baird et al., 2006). Participants were categorized as being language impaired if they had a Clinical Evaluation of Language Fundamentals—Third Edition UK (CELF-3UK; Semel, Wiig, & Secord, 2000) receptive, expressive, or total language standard
score of 77 or below, and a Wechsler Intelligence Scale for Children—III (WISC-III; Wechsler, 1992) Performance IQ or the Perceptual Organizational Index standard score of 80 or above. Sixteen adolescents with ALI (9 with a diagnosis of autism and 7 with another ASD diagnosis) and 10 with SLI were recruited from the SNAP cohort. The participants’ language and nonverbal abilities assessed for SNAP, used to establish the groups for this study, were confirmed by retesting using selected subtests from the CELF-3UK (concepts and directions [CD] and recalling sentences [RS]) and the WISC-III (picture arrangement [PA] and block design [BD]). Participants were required to have CELF-3UK CD and/or RS scaled scores below 5 and WISC-III PA and BD scaled scores above 6. Participants with language impairment were also matched on British Picture Vocabulary (BPVS; Dunn et al., 1997) scores.

In order to increase numbers in the SLI group, three additional participants with SLI were recruited from outside the SNAP cohort, from special schools for children with language impairment known to clinical services at Guy’s Hospital, London. It was not possible to complete the entire test battery of full WISC-III, CELF-3UK, ADOS-G, and ADI-R for the additional participants. The ASD status of the adolescents with SLI was assessed using the ADOS-G and Social Communication Questionnaire (Rutter, Bailey, & Lord, 2003). All three participants scored below ASD cutoffs for both instruments.

Seventeen adolescents with typical language development (TLD) matched on chronological age with the clinical groups were recruited from a single school in Southwest London. The school distributed information about the study and consent forms with prepaid envelopes to the parents of all 14-year-olds, and participants were selected from those who completed a consent form. The language and nonverbal learning abilities of the participants were screened to ensure language and nonverbal skills were in the average range using the CD and RS subtests from the CELF-3UK and PA and BD subtests from the WISC-III. The Social Communications Questionnaire was used as an autism screening measure, with no participant obtaining a score greater than 6. Each participant was offered a small cash sum to recompense their time and effort.

Table 1 shows the mean (standard deviation) standardized scores the CELF-3UK and WISC-III subtests and BPVS, together with ages and sex ratios. As expected, a series of univariate analyses of variance (ANOVA) and post hoc Tukey honestly significant difference tests indicated that the language impairment participants had weaker language skills and similar nonverbal skills compared to the normal language group, confirming their language impairment status (all \( p < .001 \)). Participants with TLD were significantly younger than those with SLI, but the three groups were matched on PA and BD scores. Individuals with SLI and ALI showed lower CD, RS, and BPVS standard scores than those with TLD.

**Design**

The gating task (Grosjean, 1980), in which listeners are presented with increasingly longer fragments of a word and are required to guess the word they are hearing, is an effective index of lexical access in real time (Tyler & Wessels, 1985). The design of the experiment was based on the cohort model of spoken word recognition
Table 1. Mean (SD) age, mean (SD) CELF-3UK subtest scores, and mean (SD) WISC-III subtest scores for the three participant groups

<table>
<thead>
<tr>
<th></th>
<th>TLD (N = 17, 7 F)</th>
<th>SLI (N = 13, 0 F)</th>
<th>ALI (N = 16, 0 F)</th>
<th>Group Differences**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (months)</td>
<td>172.7 (4.2)</td>
<td>184.2 (7.3)</td>
<td>176.3 (5.8)</td>
<td>SLI &gt; TLD = ALI</td>
</tr>
<tr>
<td>CELF-3UK</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td>10.1 (2.6)</td>
<td>4.5 (1.5)</td>
<td>4.5 (1.4)</td>
<td>TLD &gt; SLI = ALI</td>
</tr>
<tr>
<td>RS</td>
<td>9.1 (1.9)</td>
<td>3.7 (1.2)</td>
<td>4.9 (1.6)</td>
<td>TLD &gt; SLI = ALI</td>
</tr>
<tr>
<td>BPVS</td>
<td>106.4 (20.0)</td>
<td>84.2 (6.8)</td>
<td>80.4 (9.0)</td>
<td>TLD &gt; SLI = ALI</td>
</tr>
<tr>
<td>WISC-III</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>12.7 (3.9)</td>
<td>13.2 (2.6)</td>
<td>12.6 (4.0)</td>
<td>TLD = SLI = ALI</td>
</tr>
<tr>
<td>BD</td>
<td>9.6 (2.9)</td>
<td>10.2 (3.1)</td>
<td>10.6 (3.3)</td>
<td>TLD = SLI = ALI</td>
</tr>
</tbody>
</table>

Note: CELF-3UK, Clinical Evaluation for Language Fundamentals—Third Edition, UK; WISC-III, Wechsler Intelligence Scale for Children—Third Edition; TLD, typical language development; F, female; SLI, specific language impairment; ALI, autism spectrum disorder with language impairment; CD, concepts and directions; RS, recalling sentences; BPVS, British Picture Vocabulary Scales; PA, picture arrangement; BD, block design.

**p < .001.

(Marslen-Wilson, 1987). Spoken word recognition is affected by word frequency and the density of onset competitors (i.e., those sharing the same initial consonant and vowel) a word has (for an overview see Grosjean, 1996). These factors reflect the way words are stored and organized in the typically developing lexicon. By manipulating word frequency and onset density in a gating task it was possible to investigate whether lexical representation and access are typical in individuals with SLI and ALI. Word frequency (high and low) and onset density (high and low) were manipulated to create four experimental conditions: high-frequency/high-onset density, high-frequency/low-onset density, low-frequency/high-onset density, low-frequency/low-onset density (see Table 2).

**Materials**

The CELEX database (Burnage, 1990) was used to select high- and low-frequency monosyllabic words from cohorts with high- and low-onset density. Each condition consisted of eight consonant[1]vowel/consonant[1–2] words with initial consonants all either voiced or voiceless plosives. There were equal number of the voiced and voiceless items in all but the low-frequency/high-onset density condition (two voiced and six voiceless), where onset density considerations did not allow for balance. There were two word-final consonants clusters in each of the high-frequency/dense and frequency/sparse condition and one in each of the low-frequency dense and low-frequency/sparse conditions; the other items ended in a single consonant. Final consonants were more mixed, but different classes of sounds were matched across conditions. Items were matched across each variable for frequency and onset density and did not differ in duration, F (1, 3) = 0.993,
Table 2. Conditions and stimulus properties for gating experiment, including mean (SD) stimulus length

<table>
<thead>
<tr>
<th>Onset Density</th>
<th>Low&lt;sup&gt;a&lt;/sup&gt;</th>
<th>High&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low frequency</td>
<td>Dart</td>
<td>Beak</td>
</tr>
<tr>
<td>Example</td>
<td>118.8</td>
<td>106.6</td>
</tr>
<tr>
<td>Frequency</td>
<td>53.5</td>
<td>172.5</td>
</tr>
<tr>
<td>No. of competitors</td>
<td>456.7 (77.6)</td>
<td>393.8 (77.6)</td>
</tr>
<tr>
<td>High frequency</td>
<td>Gate</td>
<td>Back</td>
</tr>
<tr>
<td>Example</td>
<td>1784.0</td>
<td>1621.6</td>
</tr>
<tr>
<td>Frequency</td>
<td>60.4</td>
<td>223.5</td>
</tr>
<tr>
<td>No. of competitors</td>
<td>425.0 (65.5)</td>
<td>431.3 (70.4)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Few competitors.
<sup>b</sup>Many competitors.

$p = .410$ (see Table 2 and Appendix A). The stimuli were all recorded in a soundproof booth by a female native speaker of Southern British English. Stimuli were recorded digitally to Sony Minidisk at a sampling rate of 44.1 kHz. The recorded stimuli were imported onto a speech editing program (Audacity version 1.2.4) and split into individual sound files. The onsets and offsets of the words were determined by inspecting both the waveform, and the spectrogram. Once the onset was identified the sound files were divided into gates increasing progressively by 50 ms, for example, 50, 100, and 150 ms. This process was conducted using a macro written for Praat (Boersma & Weenink, 2011), which is available from the second author on request. The final gate was always a multiple of 50, and therefore, the offset of the word frequently occurred prior to the end. Some word-final phonemes were truncated where it was judged that enough acoustic information was present to enable their identification, and where an additional gate would contain very little information (<20 ms of signal). This procedure mainly affected sonorants. Once the gates had been selected they were examined by a second rater to check that onsets were accurate, and the final gate had been determined according to the principles outlined above.

To control for possible order effects, four different random orders of the stimuli were created. Participants within each diagnostic group were sequentially assigned to each order.

Procedure

The experimenter began with a visual analog of the gating task and two warmup items, which were administered using DMDX experimental software (Forster, 2004). First of all, two pictures were presented by revealing increasingly large segments from the bottom up. The participant was asked to guess what the picture
was at each stage until the whole picture was finally revealed. Then the experimenter said “Now we are going to do the same thing with words.” The words bed and cost were presented in gates that increased progressively by 50 ms, in accordance with the experimental words. The sound files were presented over headphones. The warmup items were repeated if the children had any difficulties understanding the paradigm.

For the experiment itself, the stimuli were presented in successive format, with all segments of each stimulus presented starting with the shortest and ending with the longest. Stimuli were presented using Microsoft Excel, with participants’ responses recorded orthographically on the spreadsheet. Nonwords were recorded pseudophonemically, using standard English orthography, for example, /peIk/ was recorded as “pake.” Sometimes a participant could not think of responses, or was reluctant to provide responses, especially at the low gates. In such cases, the experimenter attempted to cajole a response, for example, “just say the first thing that comes into your head.” If this did not work, the participant was invited to produce the first sound of the word only.

**Identification points**

Three different identification points were established for the participants’ responses, following Mainela-Arnold et al. (2008). The phoneme identification point was the gate at which the participant first identified the correct phoneme. The isolation point was the gate at which the participant first identified the word. Responses that included the word as a morpheme, for example, catching where the target was catch, were scored as correct. Total acceptance point was the gate at which the participant identified the word, and did not subsequently change their mind.

**Data cleaning and reliability**

Where participants failed to identify the word, the values for the isolation point and total acceptance point were set at the maximum gate plus 50 ms, following the approach of Walley, Michela, and Wood (1995). The gate at which children recognized the initial phoneme was also computed, and here also, where the phoneme was not recognized the gate was set at the maximum gate plus 50 ms. Outliers were identified by visual inspection of boxplots for each condition by group followed by Winsorization where outliers were replaced with Tukey’s hinges values for each condition by group. This affected 5.7% of the data. The reliability of the coding scheme was checked by a second rater. Twenty-five percent of the transcriptions were checked by an independent coder to ensure the three different recognition points for each item were accurately coded. Point-to-point reliability was high at 95.3% for the phoneme identification point, 96.6% for the isolation point, and 95.1% for the total acceptance point.

This study was approved by the South East Multicentre Research Ethics Committee (00/01/50) and De Montfort University Research Ethics Committee. Parents of the participants were asked to discuss the study with their children and signed informed consent prior to their children’s participation.
Table 3. Mean (SD) isolation points (ms)

<table>
<thead>
<tr>
<th></th>
<th>Onset Density</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>329.4 (21.4)</td>
</tr>
<tr>
<td>High</td>
<td>340.9 (24.3)</td>
</tr>
<tr>
<td>ALI</td>
<td>362.1 (15.4)</td>
</tr>
</tbody>
</table>

Note: TLD, typical language development; SLI, specific language impairment; ALI, autism spectrum disorder with language impairment.

RESULTS

Phoneme identification point

Children in all groups were quick to recognize the first phoneme of the word, which was frequently identified at the very first 50-ms gate. TLD participants correctly identified the first phoneme for 90% of words, compared with 89% in those SLI and 88% in those with ALI. The mean phoneme identification gate was 69.2 (SD = 13.8) for the participants with TLD, 63.1 (7.8), for those with SLI, and 65.4 (12.0) for those with ALI. A one-way ANOVA was conducted with diagnostic group (TLD, SLI, ALI) as the independent variable, and phoneme identification gate as the dependent variable. There was no significant effect of group, $F(2, 45) = 1.05, p > .1$, partial $\eta^2 = 0.023$.

Isolation point

The isolation point data were modeled in a mixed $3 \times 2 \times 2$ ANOVA with one between-subjects factor, diagnostic group (TLD, SLI, ALI) and two within-subjects factors, onset density (high, low) and word frequency (low, high). Table 3 shows amount of input required for participants to identify the correct word, without necessarily continuing to select the target as they hear more of the word. The data are graphically represented in Figure 1.

There was a main effect of frequency, $F(1, 43) = 83.32, p < .001$, partial $\eta^2 = 0.66$. At the isolation point high-frequency words were recognized with less input (estimated marginal mean = 289.8 ms, $SE = 2.8$ ms) than low-frequency words ($M = 322.5$ ms, $SE = 2.3$ ms). The main effect of density was not significant, $F(1, 43) = 0.04, p = .851$, partial $\eta^2 = 0.001$. The main effect of group was not significant, $F(2, 43) = 2.75, p = .075$, partial $\eta^2 = 0.113$. There was a significant crossover interaction between onset density and frequency, $F(1, 43) = 258.752, p < .001$, partial $\eta^2 = 0.857$. An analysis of simple effects indicated significant differences between low-frequency and high-frequency words in cohorts with low-onset density, $F(1, 43) = 249.89, p < .001$, partial $\eta^2 = 0.853$, and cohorts with high-onset density, $F(1, 43) = 7.16, p = .01$, partial $\eta^2 = 0.143$. Sidak corrected
Figure 1. Mean gates for isolation point by group and condition with standard error bars. OD, onset density; WF, word frequency.
contrasts showed that in cohorts with low-onset density high-frequency words were identified with less input ($M = 267.6$ ms, $SE = 4.0$ ms) than low-frequency words ($M = 344.1$ ms, $SE = 3.0$ ms, $p < .001$), whereas the reverse was true in cohorts with high-onset density: low-frequency words were recognized with less input ($M = 300.9$ ms, $SE = 3.4$ ms) than high-frequency words ($M = 312.0$ ms, $SE = 2.8$ ms, $p < .001$). The interaction between onset density and diagnostic group was significant, $F(2, 43) = 5.056, p = .011$, partial $\eta^2 = 0.19$. An analysis of simple effects indicated that there was a significant difference between the groups for words from cohorts with low-onset density, $F(2, 43) = 4.94, p = .012$, partial $\eta^2 = 0.187$, but not cohorts with high-onset density, $F(2, 43) = 2.39, p = .104$, partial $\eta^2 = 0.01$. For words from cohorts with low-onset density participants with ALI needed more input ($M = 312.7$ ms, $SE = 4.3$ ms) to identify words than those with TLD ($M = 295.0$ ms, $SE = 4.2$ ms, $p = .016$) but not those with SLI ($M = 309.9$ ms, $SE = 4.8$ ms), the difference between participants with TLD and SLI did not reach significance ($p = .071$). The interaction between frequency and group was not significant, $F(2, 43) = 2.3, p = .112$, partial $\eta^2 = 0.097$. The three-way interaction between onset density, frequency, and group was significant, $F(2, 43) = 4.88, p = .012$, partial $\eta^2 = .185$. An analysis of simple effects demonstrated this was due to differences between groups for low-frequency words from sparse cohorts, $F(2, 43) = 10.77, p < .001$, partial $\eta^2 = 0.334$. Other simple effects were not significant; high-frequency words from cohorts with low-onset density, $F(2, 43) = 1.92, p = 0.159$, partial $\eta^2 = 0.082$; low-frequency words from cohorts with high-onset density, $F(2, 43) = 0.69, p = .502$, partial $\eta^2 = 0.031$; high-frequency words from cohorts with high-onset density, $F(2, 43) = 2.55, p < .09$, partial $\eta^2 = 0.106$. In the case of low-frequency words from cohorts with low-onset density participants with ALI needed more input ($M = 362.1$ [5.1] ms) to identify words than both those with SLI ($M = 340.9$ ms, $SE = 5.7$ ms, $p = .024$) and those with TLD ($M = 329.4$ ms, $SE = 5.0$ ms, $p < .001$). Those with SLI and TLD did not differ ($p = .354$).

**Total acceptance point**

Table 4 shows amount of input required for participants to identify the correct word without subsequently changing their decision. The means are graphically represented in Figure 2.

Again, the data were analyzed in a mixed $3 \times 2 \times 2$ ANOVA with one between-subjects factor, diagnostic group (TLD, SLI, ALI) and two within-subjects factors, onset density (low, high) and word frequency (low, high). At the total acceptance point, high-frequency words were recognized with less input ($M = 351.8$ ms, $SE = 2.1$ ms) than low-frequency words ($M = 370.4$ ms, $SE = 2.0$ ms), $F(1, 43) = 87.91, p < .001$, partial $\eta^2 = 0.672$. Words from cohorts with high-onset density were recognized with less input ($M = 356.8$ ms, $SE = 2.0$ ms) than those from cohorts with low-onset density ($M = 365.4$ ms, $SE = 2.1$ ms), $F(1, 43) = 19.93, p < .001$, partial $\eta^2 = 0.317$. The main effect of group was significant, $F(2, 43) = 4.66, p = .015$, partial $\eta^2 = 0.178$. Participants with ALI needed more input ($M = 367.7$ ms, $SE = 3.0$ ms) to recognize words than those with TLD ($M = 354.6$ ms, $SE = 3.0$, $p = .012$) but not those with SLI ($M = 361.0$ ms, $SE =$
Table 4. Mean (SD) for total acceptance points (ms)

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Onset Density</th>
<th>TLD Low</th>
<th>TLD High</th>
<th>SLI Low</th>
<th>SLI High</th>
<th>ALI Low</th>
<th>ALI High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>378.1 (19.3)</td>
<td>333.4 (18.0)</td>
<td>345.2 (17.2)</td>
<td>361.8 (18.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>387.8 (16.9)</td>
<td>346.7 (7.9)</td>
<td>347.1 (16.6)</td>
<td>362.2 (6.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>406.4 (12.4)</td>
<td>339.9 (21.5)</td>
<td>357.7 (11.8)</td>
<td>366.7 (21.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: TLD, typical language development; SLI, specific language impairment; ALI, autism spectrum disorder with language impairment.

3.4, p = .387); those with SLI and TLD did not differ (p = .424). There was a crossover interaction between onset density and frequency, F (1, 43) = 298.22, p < .001, partial η² = 0.874. An analysis of simple effects indicated significant differences between low-frequency and high-frequency words in cohorts with low-onset density, F (1, 43) = 344.82, p < .001, partial η² = 0.889, and cohorts with high-onset density, F (1, 43) = 24.91, p < .001, partial η² = 0.367. In cohorts with low-onset density high-frequency words were identified with less input (M = 340.0 ms, SE = 2.6 ms) than low-frequency words (M = 390.8 ms, SE = 2.4 ms, p < .001), whereas the reverse was true in cohorts with high-onset density: low-frequency words were recognized with less input (M = 350.0 ms, SE = 2.3 ms) than high-frequency words (M = 363.5 ms, SE = 2.6 ms, p < .001). The interaction between frequency and group was significant, F (2, 43) = 6.67, p = .003, partial η² = 0.237. Analysis of simple effects revealed a significant effect of group for low-frequency words, F (2, 43) = 10.01, p < .001, partial η² = 0.318, but not high-frequency words, F (2, 43) = 1.0, p = .375, partial η² = 0.045. For low-frequency words participants with ALI needed more input (M = 382.0 [3.3] ms) to identify words than both those with SLI (M = 367.5 [3.7] ms) (p = .017) and those with TLD (M = 361.7 [3.2] ms) (p < .001). Those with SLI and TLD did not differ (p = .568). The interaction between onset density and group was not significant, F (2, 43) = 2.87, p = .068, partial η² = 0.118. The interaction between onset density, frequency, and group was not significant, F (2, 43) = 2.37, p = .105, partial η² = 0.099.

Nontarget words

An analysis of nontarget words suggested by participants during the gating task was conducted to provide an indication of the qualitative variability in responses. Table 5 shows the number of different words produced by participants.

Again the data were analyzed in a mixed 3 × 2 × 2 ANOVA with one between-subjects factor, diagnostic group (TLD, SLI, ALI) and two within-subjects factors, onset density (low, high) and word frequency (low, high). There was a main effect of frequency. Participants generated more possible nontarget words to
Figure 2. Mean gates for total acceptance point by group and frequency with standard error bars. WF, word frequency.
Applied Psycholinguistics 34:2
Loucas et al.: Spoken word recognition in ASD and SLI

Table 5. Mean (SD) number of words offered during gating

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Onset Density</th>
<th>Low</th>
<th>High</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLD</td>
<td>Low</td>
<td>3.4 (0.4)</td>
<td>3.6 (0.5)</td>
<td>3.1 (0.4)</td>
<td>3.5 (0.4)</td>
</tr>
<tr>
<td>SLI</td>
<td>Low</td>
<td>3.6 (0.4)</td>
<td>3.7 (0.4)</td>
<td>3.2 (0.4)</td>
<td>3.6 (0.4)</td>
</tr>
<tr>
<td>ALI</td>
<td>Low</td>
<td>3.7 (0.6)</td>
<td>3.8 (0.5)</td>
<td>3.4 (0.7)</td>
<td>3.8 (0.6)</td>
</tr>
</tbody>
</table>

Note: TLD, typical language development; SLI, specific language impairment; ALI, autism spectrum disorder with language impairment.

We set out to investigate the effects of word frequency and onset density on lexical access in adolescents with language impairment. These factors may affect the development of word form representations and so manipulating them offered the opportunity to probe the properties of these representations in adolescents with SLI and ALI. Our findings were consistent with previous studies of children with SLI and those with TLD (Mainela-Arnold et al., 2008; Metsala, 1998), demonstrating similar processes of spoken word recognition in adolescents with SLI and their typically developing peers, insofar as these are indexed by performance in the gating task. We found a strong interaction between frequency and density at isolation point and total acceptance point, such that high-frequency words with low-onset density
density are recognized with less input than those with many onset competitors. In contrast, low-frequency words with high-onset density are recognized with less input than those with low-onset density. Low-frequency words with low-onset density required most acoustic–phonetic input for recognition. This pattern of performance may be explained on the assumption that word form representations become precisely specified only gradually during development. Representational restructuring may be driven by repeated exposure to tokens of particular words; hence, the effect of frequency, and the pressure to distinguish between similar sounding competitors, greater in environments with high-onset density, hence the effect of neighborhood density (cf., Walley, 1993). Thus, low-frequency words with few competitors will be under the least pressure to restructure phonological representations, their representations will be least well specified and so they will require most acoustic–phonetic input to be recognized.

We extended Mainela-Arnold et al.’s (2008) study through our use of a different definition of competitor environment, onset density as opposed to neighborhoods, and the inclusion of a group with ALI in addition to individuals with SLI. The adolescents with SLI and those with TLD showed a similar interaction between word frequency and competitor density, suggesting that lexical competition defined in these different ways has similar effects in gating. We note, however, that studies that have found similar gating performance in SLI and typically developing controls, at least for familiar words, have involved children from middle to late childhood, and that deficits may be found in younger children.

To our knowledge, this is the first investigation of spoken word recognition in individuals with ASD who also have language impairment using the gating task. In many ways the adolescents with ALI showed similar spoken word recognition in the gating task when compared to the SLI and TLD groups. In general, there were notable parallels between the diagnostic groups. There was no main effect of group at the phoneme identification point, the isolation point, suggesting all three groups required similar amounts of input to make their decisions. There was an effect of group at total acceptance point, where the participants with ALI required more input than those with TLD to identify targets. This was also reflected in the ALI group generating more non-target words overall. At the total acceptance point and in the number of non-target responses the ALI group did not differ to the SLI group, which in turn, did not differ to the TLD group. The similarity between the SLI and TLD groups on these latter measures marks a difference with Mainela et al. (2008), who found their participants with SLI identified words later at the total acceptance point and generated more non-target words, and may reflect the greater age and language level of the participants here.

The manipulation of lexical variables allowed us to probe the process in a more fine-grained way and here we found an interaction between competitor density, word frequency, and diagnostic group at the isolation point and interaction between frequency and diagnostic group at the total acceptance point. In both cases it was the individuals with ALI who drove the interaction. They required more acoustic phonetic input to recognize low-frequency words in low-onset density neighborhoods than either the participants with SLI or those with TLD at isolation point and more input to recognize low-frequency words at the total acceptance point.
Although all participants regardless of diagnostic status showed effects of word frequency and onset neighborhood density, suggesting a qualitatively similar pattern of spoken word recognition, these effects were more marked in the ALI group. Insofar as efficient lexical access relies on the structure of underlying phonological representations, this may be evidence that they may be less well specified in the adolescents with ALI than those with SLI or TLD. The finding may be explained on the assumption that if constraints driving the restructuring of phonological representations are relatively weak, for instance, when words have few phonological neighbors and/or are low in frequency, deficits in the development of word forms are more apparent. The impairments in speech processing reported in children with ASD could lead to less efficient use of input to build lexical representations. Speech processing differences in ASD include atypical orienting to speech (Ceponiene et al., 2003); an atypical preference for the nonspeech signals and failure to respond to a syllable change (Kuhl et al., 2005); poor speech-in-noise perception suggesting a reduced ability to exploit speech information during the spectral and temporal dips in the background (Alcantara, Weisblatt, Moore, & Bolton, 2004); speech-evoked responses in quiet and background noise show minimal quiet-to-noise differences because quiet responses are already severely degraded (Russo, Zecker, Trommer, Chen, & Kraus, 2009). If these processing deficits impact on the ability to extract acoustic–phonetic information from the speech signal effectively then it would be more difficult to construct well-specified word form representations. In adolescents with ALI the downstream effects of this lower level processing deficit may only emerge for words that are low in frequency and/or have few onset neighbors, leading to the quantitative differences in gating performance reported here. This account would predict that earlier in development differences between children with ASD and children with SLI and TLD would be greater than reported here and emerge for word form representations more generally.

If individuals with ALI have poorly specified word form representations the effects should impact other aspects of language processing. This may be evidenced by the poor performance on standardized measures of verbal comprehension found in ASD, which is shown most directly in poor word comprehension in preschoolers (Charman, Drew, Baird, & Baird, 2003; Hudry et al., 2010) but also in older children and adolescents (Loucas et al., 2008). However, it is not possible to distinguish between deficits at different levels of language that may affect performance on broad measures of comprehension. Failure on a test may be due to deficits in representations of lexical form, but may also be the result of poor lexical semantic or syntactic knowledge, or indeed nonlinguistic factors such as attention and memory. The results we present suggest impairments at the level of word form representations and so it may be expected that other tasks tapping phonological representations should be affected by any such deficits. However, other evidence does not indicate impaired phonological representations. In nonword discrimination adolescents with ALI are able to distinguish stimuli that differ by a single phonetic feature as well as age-matched controls (Loucas et al., in press). In addition, the results presented here for identification of initial phoneme show that participants with ALI did not differ from those with TLD and SLI in the amount of acoustic–phonetic input they need to identify the phoneme accurately. However, phonological processing tasks that are not associated with particular
word forms can draw on abstract phonological representations that emerge from interactions between word form representations within the lexicon as a whole. This suggestion is consistent with Storkel and Morrissette’s (2002) model of the lexicon in which lexical and phonological representations develop interdependently. The model allows for phonological overlap at the lexical level to support the formation of representations at the phonological level and in this way provide the basis for apparently intact general phonological processing ability. But, when specific lexical representations are probed, deficits in spoken word recognition may still emerge.

Although the results presented here may be explained by a speech processing deficit, they may reflect a more general deficit in forming mental representations resulting from differences in perception and cognition in ASD. Plaisted (2001) has argued cognition in ASD is characterized by reduced generalization that results from an impaired ability to perceive similarities between stimuli. Alternatively, Mottron, Dawson, Soulèères, Hubert, and Burack (2006) argue that individuals with ASD show enhanced perceptual functioning presenting as superiority in low-level perceptual processing coupled with diminished processing of complex material. Both of these theories may allow accurate acoustic–phonetic processing but a diminished ability to form more abstract and complex representations of phonological form. Thus, in the case of reduced generalization, a diminished ability to recognize similarities between the phonological form of spoken words may result in a reduced ability to use onset overlap to support the building of representations lexical form. Or, in the case of enhanced perceptual functioning, diminished processing of complex material, such as phonological forms, may lead to less well-specified word form representations. However, as these cognitive differences are not limited to speech processing, impairments or differences in other aspects of lexical representation would be predicted, such as lexical semantics, as well as differences in the properties of representations in other cognitive domains.

CONCLUSIONS

Our results provide some preliminary evidence that adolescents with ALI may be impaired in some aspects of spoken word recognition, relative to those with SLI. Although overall there were notable similarities in the overall pattern of performance on the gating task reported here, the participants with ALI showed greater effects of onset density and word frequency than those with SLI for low-frequency monosyllables with few competitors. The SLI group did not differ with the TLD group for these stimuli. Although others have also found similar gating performance in children with SLI compared with those with TLD, the detail of these studies does indicate some differences in SLI. Dollaghan (1998) found that children with SLI required more input than typically developing children to recognize newly learned words, and Mainela-Arnold et al. (2008) found some evidence that children with SLI were uncertain about their lexical choices at late gates argued, which may be the result of a difficulty forming new phonological representations. In other paradigms, such as lexical decision, children with receptive SLI may show clearer evidence of impaired phonological representations (Edwards & Lahey, 1996). Thus, strong conclusions about underlying differences
between SLI and ALI cannot be drawn on the basis of results from a particular experimental paradigm. However, the results presented here invite further research in this area. If these findings are replicable and emerge in different paradigms the issue of the underlying cause will need to be addressed; in particular, whether the apparently subtle deficits reported here that emerge in difficult-to-form lexical representations are the result of speech processing deficits.

Other studies of language impairment in ALI and SLI are equivocal as to whether the pattern of linguistic deficits is similar or different. Similarities include evidence for intact speech perception and phonological short-term memory processes (Loucas et al., 2010), and deficits in inflectional morphology (Roberts, Rice, & Tager-Flusberg, 2004) and in using context to disambiguate word meanings (Norbury, 2005). Differences include more severe comprehension problems in ALI than SLI (Loucas et al., 2008), but greater syntactic difficulties in sentence repetition in SLI than ALI (Riches, Loucas, Baird, Charman, & Simonoff, 2010). These results on spoken word recognition may point to a further difference in the underlying deficits, with ALI characterized by less well-specified word form representations affecting spoken word recognition.
APPENDIX A

Word frequency (WF) and onset density (OD; number of cohort competitors) for four categories of target words

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Boost</td>
<td>149</td>
<td>66</td>
<td>Ball</td>
<td>1996</td>
<td>84</td>
<td>Beak</td>
<td>120</td>
<td>99</td>
<td>Back</td>
<td>1324</td>
<td>345</td>
</tr>
<tr>
<td>Goose</td>
<td>188</td>
<td>19</td>
<td>Birth</td>
<td>1128</td>
<td>95</td>
<td>Cub</td>
<td>76</td>
<td>290</td>
<td>Base</td>
<td>1432</td>
<td>96</td>
</tr>
<tr>
<td>Kite</td>
<td>83</td>
<td>26</td>
<td>Cook</td>
<td>1284</td>
<td>30</td>
<td>Dot</td>
<td>170</td>
<td>136</td>
<td>Boat</td>
<td>1368</td>
<td>98</td>
</tr>
<tr>
<td>Kerb</td>
<td>107</td>
<td>66</td>
<td>Dance</td>
<td>1177</td>
<td>41</td>
<td>Paste</td>
<td>92</td>
<td>143</td>
<td>Card</td>
<td>1251</td>
<td>160</td>
</tr>
<tr>
<td>Pike</td>
<td>114</td>
<td>86</td>
<td>Gate</td>
<td>1231</td>
<td>65</td>
<td>Peep</td>
<td>64</td>
<td>138</td>
<td>Catch</td>
<td>3459</td>
<td>498</td>
</tr>
<tr>
<td>Toad</td>
<td>76</td>
<td>78</td>
<td>Task</td>
<td>1484</td>
<td>48</td>
<td>Pod</td>
<td>63</td>
<td>229</td>
<td>Desk</td>
<td>1633</td>
<td>280</td>
</tr>
<tr>
<td>Gorge</td>
<td>142</td>
<td>42</td>
<td>Touch</td>
<td>1967</td>
<td>92</td>
<td>Purse</td>
<td>185</td>
<td>170</td>
<td>Kid</td>
<td>1411</td>
<td>125</td>
</tr>
<tr>
<td>Dorm</td>
<td>91</td>
<td>45</td>
<td>Town</td>
<td>4005</td>
<td>28</td>
<td>Putt</td>
<td>83</td>
<td>175</td>
<td>Path</td>
<td>1095</td>
<td>186</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>118.75</td>
<td>53.5</td>
<td>1784</td>
<td>60.375</td>
<td>106.625</td>
<td>172.5</td>
<td>1621.625</td>
<td>223.5</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS
The authors thank the parents/guardians and individuals who participated in the study; Autism Speaks/The National Alliance for Autism Research for their generous funding; and Abigail Davison-Jenkins, Ann Ozsivadjian, and Vicky Slonims for their help in screening the participants with autism spectrum disorders. The authors also thank the two anonymous reviewers for their valuable comments on the manuscript.

REFERENCES


