Language after childhood hemispherectomy: A systematic review

Andrea S Nahum, MSci; Frédérique J Liégeois, PhD

Author Affiliations: Cognitive Neuroscience and Neuropsychiatry Section, UCL Great Ormond Street Institute of Child Health, London, UK (*Nahum, Liégeois*); and UCL Medical School, University College London, London, UK (*Nahum*)

Corresponding Author:
Dr Frédérique Liégeois (f.liegeois@ucl.ac.uk)

Submission type: Views & Reviews

Word Count:

- *Title:* 59 characters (including spaces)
- *Abstract: 244*
- *Main Text: 4,288*
- *Figures*: 5
- *Tables:* 3
- *References*: 60

Supplementary data & material: PRISMA Checklist

Publication History: None

Statistical/ Data analysis: Andrea Nahum under the supervision of Dr Frédérique Liégeois (see affiliations above)

Search terms:

- [66] Epilepsy Surgery
- [200] Aphasia
- [205] Neuropsychological assessment
- [227] All Pediatric

Study funding: No targeted funding reported.

Disclosures: Andrea Nahum and Dr Frédérique Liégeois report no disclosures relevant to the manuscript.

Abstract

Objective

To conduct a systematic review on language outcomes after left and right hemispherectomy in childhood, a surgical procedure that involves removing or disconnecting a cerebral hemisphere.

Methods

We searched MEDLINE, Embase and PsycInfo for articles published between 1st January 1988 and 16th May 2019. We included: (1) all types of observational studies; (2) studies where hemispherectomy was carried out before age 18 and; (3) studies with standardised scores measuring receptive vocabulary, expressive vocabulary, sentence comprehension, and/or sentence production. We calculated mean z-scores after left and right hemispherectomy in the whole group and within aetiology-specific subgroups.

Results

Our search identified 1096 studies, of which 17 were eligible. The cohort added up to 205 individuals (62% left hemispherectomy) assessed one to 15 years after surgery. In the left surgery group, all language skills were impaired (z scores <-1.5) except sentence comprehension. In the right surgery group, language performance was in the borderline range (z scores ~ -1.5). Children with cortical dysplasia showed the worst outcomes irrespective of surgery side (z scores <-2.5). Individuals with left vascular aetiology and right-sided Rasmussen Syndrome showed the best outcomes.

Conclusion

Evidence based on the largest patient cohort to date (205 participants) suggests that the risk of language impairment after hemispherectomy is high, with few exceptions. Aetiology plays a major role in post-surgical plasticity. We recommend specialist evaluation of language skills soon after surgery to identify intervention targets. Large scale studies examining outcomes in consecutive cases are still needed.

Glossary

AEDs = anti-epilpeptic drugs; **CD** = Cortical Dysplasia; **CI** = Confidence Interval; **EBNP** = evidence-based neuropsychology checklist; **HME** = Hemimegalencephaly; **I** = infarct; **LX** = left hemispherectomy; **PRISMA** = Preferred Reporting Items for Systematic Reviews and Meta-Analysis; **RS** = Rasmussen Syndrome; **RX** = right hemispherectomy; **SW** = Sturge-Weber Syndrome; **WNL** = within normal limits.

Epilepsy is a common disorder among children with a worldwide prevalence ranging from 50 to 100/100,000.[1] Antiepileptic drugs usually enable seizure control, but in up to a third of cases seizures are refractory to medical therapy.[2] Surgical removal of the epileptogenic area can be an effective treatment in selected cases.[2] For drug-resistant epilepsy due to unilateral pathology, hemispherectomy (removal or disconnection of an entire cerebral hemisphere) leads to seizure freedom in over 60% of patients.[3,4] Hemispherectomy is indicated for conditions such as Rasmussen Syndrome (an auto-immune encephalopathy with postnatal onset), Sturge-Weber Syndrome (vascular), hemimegalencephaly, cortical dysplasia and vascular insults.[5] It is primarily performed in children due to their supposed greater capacity for plasticity and functional reorganisation.[2,6]

Cognitive functions have been reported to improve substantially after hemispherectomy although they can vary from "normal" (age-appropriate) to severely impaired.[7,8] Case studies have suggested that expressive and receptive language functions can be completely spared after left hemispherectomy.[9,10] This positive outcome supports the idea of functional plasticity in the child brain and complements functional MRI studies showing that language functions can "shift" to the right hemisphere.[11–14] Yet language may be less proficient in left than right hemispherectomy patients, especially when disease onset is postnatal.[6] The literature on language outcomes comprises mainly small sample size studies that need to be regrouped and reviewed to guide clinical decisions and improve prognosis.

We conducted a systematic review of published studies to (i) report language performance in children after left versus right hemispherectomy and (ii) explore language outcomes by aetiology.

Methods

Search Strategy and Selection Criteria

We conducted a systematic literature search to identify relevant studies. We searched MEDLINE, Embase and PsycInfo for papers published between 1st January 1988 and 16th May 2019 with the term "language" combined with the terms "hemispherectomy" [MeSH] "hemispherotomy", "hemidisconnection" "hemidecortication", "hemicorticectomy" "hemispheric surgery". No language restrictions were applied. We manually screened the references of identified articles for additional relevant studies. Inclusion criteria were: (1) all types of observational studies, including single case reports; (2) studies where hemispherectomy was carried out before age 18 and; (3) studies with relevant language specific data (receptive and/or expressive vocabulary and sentence comprehension and/or production). We excluded non-peer-reviewed articles, conference abstracts, reviews, books, animal and adult studies.

Data analysis

Two independent reviewers (A.S.N. and F.J.L.) examined titles, abstracts, and studies using the criteria mentioned above. They resolved disagreements through discussion. We extracted sample size, sex, aetiology, age at seizure onset, age at surgery and age at testing, as well as side of resection and seizure outcome. We excluded studies mentioning duplicate data from other papers or marked them explicitly as having duplicate patients' data. We did not include conference proceedings as these rarely provide individual scores.

We examined four post-surgical language measures: receptive vocabulary, expressive vocabulary, sentence comprehension, and sentence production. When several language assessment scores were reported at different ages, we used the latest post-surgical assessment. Language data outside our focus, assessing measures such as phonology, word repetition, fluency, mean length of utterance, lexical judgement, syntactic judgement, written language, and verbal IQ, was excluded from our analysis. We also excluded studies providing an overall language assessment, or a more general expressive language and /or receptive language score.

Data extraction

To allow comparison between studies, we did not include data in the review if they were reported as age equivalent, percent correct, raw scores, or percentiles. We used z-scores from standard scores as outcome measures. When z-scores were not already provided in articles, we used the classical formula to calculate them based on test norms. Standardised language tests (e.g. PPVT, BPVS) have known means and standard deviations based on population data. One study did not use a standardised assessment and instead provided the means and standard deviation from age matched "healthy normal volunteers". We therefore calculated language 'z-scores' based on this group's data for this study.

To obtain overall means, as well as 'by study' means, we averaged single case values and combined means from multiple sub-groups and different studies using the weighted arithmetic mean.[15] We used the pooled standard deviation formula to estimate variation and calculate 95% confidence intervals (CI).

For interpretation, we defined impairment as a score of -1.5 SD below the mean or lower, as used in most clinical settings.[16] For scores above this threshold, performance was defined as "within normal limits" (WNL). When three or more studies reported a given language measure, we presented the data in forest plots made using the DistillerSR Forest Plot Generator from Evidence Partners.[17] For multi-subject studies, we show the data spread based on 95% CI except for one study for which no estimate of variation was provided.

When data were available for groups of 10 patients or more, we compared patient means to those of the general population (z score=0, SD=1) using one-sample t-tests with unknown standard deviation.

Risk of bias

Both authors examined selected articles independently for risk of bias within studies, namely blinding of outcome collectors to side and aetiology, and potential patient selection.[18] Risk of bias across studies was estimated using funnel plots examining effect size as a function of sample size, created with IBM SPSS Statistics Version 26.0, where 10 studies or more were available (i.e. for receptive vocabulary in the whole left surgery group).

Quality appraisal

We assessed the quality of each study using a modified version of the 19-item evidence-based neuropsychology checklist (EBNP), specifically designed for evaluating neuropsychological outcomes after epilepsy surgery.[19] A score of either 0 (Y=yes) or 1 (N=no) was given to each question and a total score was calculated. We classified studies in three groups: the high-quality group with total scores of seven, the medium quality group with total scores between five and six, included, and a low-quality group with total scores of four or below.

Meta-analysis

Where individual data were available, we ran Pearson and Spearman's rank correlations to examine the relationship between language outcome and (i) age at surgery, and (ii) age at onset of epilepsy, in exploratory analyses.

Data Availability

Data are available to qualified investigators on request to the corresponding author.

Results

Our search identified 1096 publications, excluding duplicates. From the titles, we selected 365 for review of the abstract and finally 223 for detailed review (Figure 1). We excluded 206 publications after assessment of full text because they did not meet the inclusion criteria. The main reasons for exclusion were lack of standardised language scores (e.g. development quotient or raw score), epilepsy surgery studies with no separate hemispherectomy data, and inadequate article type (e.g. review, comments, letters). Seventeen articles, all in English, were included in our systematic review.[6,7,10,13,14,20–31]

Eleven studies were outcome studies with language data and six were neuroimaging studies with language data. Seven studies had sample sizes under five (four were single case studies) and three studies had sample sizes greater than 30.

The seventeen selected studies (see Table 1 for details) contained data from a total of 205 patients, of which 77 (38%) had undergone right hemispherectomy and 128 (62%) left hemispherectomy. Mean age at seizure onset was 3.9 ± 3.5 years (range: 0 to 13.5 years) and mean age at surgery was 7.5 ± 4.7 years (range: 0.3 to 16.1 years). Mean time from onset of seizures to surgery was 3.8 ± 3.3 years (range: 0.3 to 14.3 years). Mean age at testing was 13.7 ± 5.8 years (range: 7.3 to 22.8 years) and mean time from surgery to testing was 6.1 ± 4.3 years (range: 0.9 to 15 years). Excluding one study (n=64) for which seizure freedom data were not extractable, 140 out of 141 patients (99%) were seizure free.[7]

Quantitative data were available for all four language measures, although the tests used varied between studies (Table 2).[32–40]

Language outcomes – All aetiologies confounded

Fourteen studies reported receptive vocabulary scores. In the left hemispherectomy group, mean scores were in the impaired range (-1.5 SD below the mean or lower) in seven studies (n=54) and within normal limits (WNL) in seven studies (n=43, see Figure 2A). In the right hemispherectomy group, mean scores were in the impaired range in four studies (n=54) and WNL in three studies (n=7, see Figure 2B).

Nine studies reported expressive vocabulary scores. In the left hemispherectomy group, mean scores were in the impaired range in five studies (n=32) and WNL in four studies (n=9, see Figure 2C). In the right hemispherectomy group, mean scores were in the impaired range in one study (n=3) and WNL in three studies (n=35, see Figure 2D).

Six studies reported sentence comprehension scores. In the left hemispherectomy group, mean scores were in the impaired range in one study (n=1) and WNL in five studies (33 patients; see Figure 2E). After right hemispherectomy, the mean z-score for sentence comprehension was -1.27 (n=17; CI=-0.66; -1.87), based on two studies: Save-Pédebos et al. (z-score=-1.44; n=16; CI=-1.93; -0.95) and Villarejo-Ortega et al. (z score=1.50; n=1).[24,26]

Two studies reported sentence production scores after left hemispherectomy. The mean z-score was -1.73 (n=7; CI=-2.75; -0.71) based on two studies: Bulteau et al. (z-score=-2.25; n=6) and Grosmaitre et al. (z score=1.39; n=1).[10,28] No sentence production scores were reported after right hemispherectomy.

Effect of Age at Surgery

There was no correlation between age at surgery and receptive vocabulary for the whole group (left cases: n=51, r=0.19, p=0.172; right cases: n=13, rho=0.35, p=0.238) or for any subgroup (p>0.12). Similarly, there was no correlation between age at surgery and expressive vocabulary in the left group. There was a positive correlation for right surgery cases (n=5, rho=0.90, p=0.037), and all had Rasmussen syndrome. Numbers were too small (n<5) to examine the effect of age at surgery on sentence comprehension and production.

Effect of Age at Seizure Onset

Overall, there was no correlation between age at seizure onset and receptive vocabulary for the whole group (left cases: n=50, rho=0.094, p=0.518; right cases: n=13, r=0.414, p=0.160). There was no correlation between age at seizure onset and expressive vocabulary, but numbers were small (left: n=8, rho=-0.325, p=0.432; right: n=5, rho=0.700, p=0.188).

Numbers were too small to examine the effect of seizure onset on sentence comprehension (LX=4; RX=1) and production (LX=1; RX=0).

Comparison with the general population

All language scores were lower in left hemispherectomy patients than in the general population, with mainly large effect sizes (Receptive vocabulary: t(96)=-12.42, p<0.001, d= -1.26; 95% CI= -1.49; -1.03; Expressive vocabulary: t(40)=-9.11, p<0.001, d= -1.42; 95% CI= -1.80; -1.05; Sentence comprehension: t(33)=-4.67, p<0.001, d= -0.80; 95% CI= -1.13; -0.47).

All language scores were also lower in right hemispherectomy patients than in the general population, with mainly large effect sizes (Receptive vocabulary: t(60)=-8.02, p<0.001; d= -1.03;

95% CI= -1.29; -0.76; Expressive vocabulary: t(37)=-4.90, p<0.001, d= -0.79; 95% CI= -1.11; -0.48; Sentence comprehension: t(16)=-4.33, p<0.001; d= -1.05; 95% CI= -1.58; -0.52).

Language outcomes – Rasmussen Syndrome

Eight studies reported receptive vocabulary scores. In the left hemispherectomy group, mean scores were in the impaired range in four studies (n=19) and WNL in four studies (n=13, see Figure 3A). In the right hemispherectomy group, mean scores were WNL in all five studies (n=29, see Figure 3B).

Six studies reported expressive vocabulary scores. In the left hemispherectomy group, mean scores were in the impaired range in three studies (n=18) and WNL in three studies (n=8, see Figure 3C). In the right hemispherectomy group, mean scores were WNL in three studies (n=22, see Figure 3D).

Four studies reported sentence comprehension scores. In the left hemispherectomy group, mean scores were in the impaired range in one study (n=1) and WNL in three studies (n=8, see Figure 3E). After right hemispherectomy, only Villarejo-Ortega et al. reported sentence comprehension (z score=1.50; n=1).[26]

Two studies reported sentence production scores. The mean z-score was -1.73 (n=7; CI=-2.75; -0.71) based on Bulteau et al. (z-score=-2.25; n=6) and Grosmaitre et al. (z score=1.39; n=1), for the left hemispherectomy group.[10,28] No sentence production standard scores were reported after right hemispherectomy.

Effect of Age at Surgery

Age at surgery did not correlate with receptive vocabulary in the left (n=20, r=-0.048, p=0.841) or right (n=11, r=0.473, p=0.142) surgery groups. There was no significant correlation between age at surgery and expressive vocabulary for left cases (n=7, rho=-0.259, p=0.574), but there was for right cases, although numbers were small (n=5, rho=0.900, p=0.037).

Effect of Age at Seizure Onset

There was no significant correlation between receptive vocabulary and age at seizure onset for left (n=20, r=0.022, p=0.928) or right (n=11, r=0.396, p=0.227) surgery cases. Similarly, there was no significant correlation between expressive vocabulary and age at seizure onset for left (n=7, rho=-0.308, p=0.502) or right (n=5, rho=0.700, p=0.188) surgery cases.

Comparison with the general population

For the left surgery group, all language scores were lower than those of the general population, with large effect sizes (Receptive vocabulary: t(35)=-10.85, p<0.001; d= -1.81; 95% CI= -2.26; -1.35; Expressive vocabulary: t(28)=-7.10, p<0.001, d= -1.32; 95% CI= -1.75; -0.89).

For the right surgery group, language scores were also lower than those of the general population, with medium effect sizes (Receptive vocabulary: t(31)=-3.54, p<0.001, d= -0.63; 95% CI= -0.95; -0.30; Expressive vocabulary: (t(24)=-3.65, p<0.001, d= -0.73; 95% CI= -1.11; -0.34).

Sample sizes were too small to examine sentence comprehension and production scores.

Language outcomes – Vascular Aetiologies

Eight studies reported receptive vocabulary scores. In the left hemispherectomy group, mean scores were in the impaired range in two studies (n=3) and WNL in six studies (n=27, see Figure 4A). After right hemispherectomy, only Pulsifer et al. reported receptive vocabulary (z score=-2.07; n=5; CI=-2.60; -1.53).[7]

Two studies reported expressive vocabulary scores. In the left hemispherectomy group, the mean z-score was -2.86 (n=3; CI=-5.94; 0.21) based on Danelli et al., (z-score=-4.59; n=1) and Pulsifer et al. (z score=-2.00; n=2; CI=-6.44; 2.44).[7,30] After right hemispherectomy, only Pulsifer et al. reported expressive vocabulary (z score=-0.29; n=5; CI=-0.98; 0.39).[7]

Danelli et al. reported sentence comprehension after left hemispherectomy (z score=0.85; n=1).[30] No sentence comprehension scores were reported after right hemispherectomy.

No study reported sentence production scores.

Effect of Age at Surgery

There was no significant correlation between receptive vocabulary and age at surgery in the left surgery group (n=28, rho=0.110, p=0.578). Due to small sample sizes (n=5 or less) we were unable to examine the effect of age at surgery for the right group, or on other language skills.

Effect of Age at Seizure Onset

There was a significant correlation, with a medium effect size, between receptive vocabulary and age at seizure onset for left surgery cases (n=27, r=0.469, p=0.014). Due to lack of data, we were unable to examine this question for the right surgery group. Similarly, we were unable to examine the relationship between expressive vocabulary and age at seizure onset.

Comparison with the general population

For the left surgery group, receptive vocabulary scores were lower than those of the general population, with a large effect size ($t(29)=-6.02$, $p<0.001$, $d= -1.10$; 95% CI= -1.49; -0.71). Sample sizes were too small to compare other scores to those of the general population.

Language outcomes – Cortical Dysplasia

Four studies reported receptive vocabulary scores. In the left hemispherectomy group, mean scores were in the impaired range in two studies (n=12) and WNL in two studies (n=2, see Figure 4B). After right hemispherectomy, only Pulsifer et al. reported receptive vocabulary (z score=-2.96; n=9; CI:-3.91; -2.01).[7]

Two studies reported expressive vocabulary scores. In the left hemispherectomy group, the mean z-score was -2.87 (n=9; CI=-3.73; -2.00) based on Ibrahim et al. (z-score=-1.47; n=1) and Pulsifer et al. (z score=-3.04; n=8; CI=-3.94; -2.14).[7,21] After right hemispherectomy, only Pulsifer et al. reported expressive vocabulary (z score=-2.64; n=8; CI=-3.98; -1.30).[7]

No study reported sentence comprehension or production scores.

Effect of Age at Surgery

This question could not be examined due to small sample size (N<5).

Effect of Age at Seizure Onset

Due to lack of data (n<5 we were unable to examine this question.

Comparison with the general population

For the left surgery group, receptive vocabulary scores were lower than those of the general population, with a large effect size (t(13)=-9.76, p<0.001, d= -2.61; 95% CI= -3.60; -1.62).

Sample sizes were too small to compare other scores to those of the general population.

Risk of bias

As can be seen from Figure 2, smaller studies tended to report more favourable outcomes than both the larger study and the pooled mean. We therefore cannot rule out a bias towards reporting more favourable receptive vocabulary outcomes in the left surgery group.

No study reported whether language outcomes were measured by individuals who were blinded to side or aetiology. The risk of bias for patient selection was assessed using items 3 and 4 from our critical appraisal checklist (Table 3). We chose not to exclude any study based on risk of bias.

Quality Appraisal

Quality appraisal scores for each paper can be found in Table 3. Seven studies (41%) were of high quality, seven studies (41%) were of medium quality and three studies were of low quality (18%).

Discussion

Our systematic review showed that patients with left-sided surgery have clinical impairments in receptive vocabulary, expressive vocabulary and sentence production. The only exception was sentence comprehension, with scores at the low end of the average range. After right hemispherectomy, in contrast, all language scores fell at the low end of the typical range but no data were available for sentence production. Importantly, there were wide variations in individual scores, and outcome profiles depended on aetiology. Young people with right-sided Rasmussen Syndrome showed no impairment. Patients with left-sided vascular aetiology showed preserved sentence and single-word comprehension but impaired expressive vocabulary. Children with cortical dysplasia showed the worst outcomes, with impairments in all skills measured, irrespective of side of surgery (Figure 5).

Overall, language scores after right and left hemispherectomy were significantly lower than the population means, indicating that language skills are at risk in this population as a whole. Of note, no patient was reported to use alternative or augmentative communication systems, and functional conversation skills were reported anecdotally.[13, 20, 22, 27, 28] Similarly, a large study examining spoken language outcomes via questionnaires reported that only 30% of patients had limited language (2-3 phrases or fewer) after surgery.[8]

Based on current neuroanatomical models of language, we hypothesized relatively preserved semantic skills as measured by single-word knowledge.[41,42] As these skills are subserved by a bilateral ventral language route present from birth, we predicted that the effect of a large unilateral surgery/pathology would be compensated for by the remaining right hemisphere.[43,44] Indeed, a few previous studies had reported relatively preserved single-word comprehension relative to grammatical comprehension, and one even reported worse expressive vocabulary in right than left hemispherectomy.[6,45]

Other studies suggested no effect of side of surgery on receptive vocabulary.[46] Our data, derived from the largest sample to date, indicate a persistent risk of both receptive and expressive vocabulary impairment after left and right hemispherectomy. Given the lack of evidence based interventions for children with extensive brain injury, children who undergo hemispherectomy (except for right-sided Rasmussen Syndrome) may benefit from vocabulary interventions known to be efficient for developmental language disorders.[47] Addressing this vocabulary gap may, in turn, help support literacy development.[48] Importantly, the children's apparent ability to follow conversations in everyday settings may mask substantial language needs.

In the group with Rasmussen Syndrome, we observed the expected preserved single-word skills after right-sided pathology, consistent with studies that examined postnatal/late onset pathology separately.[6,25] Given that Rasmussen Syndrome is a progressive disease manifesting, at the earliest, in the second year of life, our data show evidence of a left hemisphere advantage for single-word knowledge after the age of two. Similarly, other groups have reported language impairments post stroke after the age of two, although partial compensatory potential of the right hemisphere may subside until age five.[49]

In the vascular aetiology group, our data clearly highlight the need for long-term language support after hemispherectomy. The group with cortical dysplasia showed a distinct profile, with all language functions severely impaired irrespective of hemispheric side of surgery. There are two possible interpretations for this finding that are not mutually exclusive. Firstly, anomalies of cortical development such as cortical dysplasia may be only partially unilateral and could affect both hemispheres, with surgery performed on the most affected side.[50] Recently developed advanced automatic MRI methods may improve detection of these anomalies.[51] Secondly, hemispherectomy for cortical dysplasia results in poorer seizure control than for other pathologies, which in turn may result in worst cognitive and language outcomes.[52]

The poor language outcomes of both perinatal stroke and cortical dysplasia patients with right-sided pathology support a crucial early role of the right hemisphere for the development of language functions, as suggested by previous reports.[14,53] This early right hemisphere involvement is consistent with fMRI studies indicating right hemisphere dominance for speech processing, and left-right functional connectivity in the posterior superior temporal gyrus at rest.[43,54,55]

In the subset of data available, we failed to identify a relationship between age at surgery or age at onset of epilepsy and vocabulary outcome. We found no evidence for a cut-off age either, beyond which left hemispherectomy results in poorer outcome. These findings would go against the traditional view of a critical period for language acquisition.[11] We suggest that receptive vocabulary skills can develop after left hemispherectomy in childhood, irrespective of age at surgery. In the group with vascular aetiology only, later age at seizure onset was associated with better outcome, confirming the disruptive effect of seizures on development previously reported in children with perinatal infarcts.[12] A noteworthy finding was that later right hemisphere surgery was associated with better post-operative expressive vocabulary in the Rasmussen group. This early vulnerability supports the idea of early right hemisphere involvement in the development of this function.[14, 53] As data were only available for a handful of cases, these results must be interpreted with caution. Future studies should examine the effect of age at surgery on a range of language skills in a larger sample.

Limitations & Recommendations

Despite reporting on a large cohort, we were unable to investigate the effect of different post-operative periods, age at surgery, and age at seizure onset, and sex on language outcome across our full sample. In left hemispherectomy patients, longer post-surgical evaluation periods have been associated with higher language performances in some studies, but not others.[56,57] Exceptionally late postsurgical language onset (9 years) has also been reported.[58] Early childhood surgery has been associated with better expressive language after right hemispherectomy in acquired pathologies.[52] However, other studies have found no correlation between age at surgery and our four language measures.[29] Altogether, only large scale multi-centre longitudinal studies would give an accurate indication of the trajectory of language abilities after hemispherectomy. The gap between what is expected for a child's age and their actual language performance may widen as they grow older, despite significant absolute gains in language skills. In other words, whether patients post hemispherectomy "grow into" language deficits or "catch up" remains unknown. Future research should synthesize the evidence from longitudinal studies to report how language skills change from pre-to postoperative status, and to examine the trajectory of language development after surgery.

We were also unable to gather extensive evidence for sentence level language skills. Overall, receptive vocabulary was the most reported language measure (14 studies), followed by expressive vocabulary (nine studies). Sentence comprehension and production measures were much scarcer (respectively six and two studies), notably for the right hemispherectomy group. Additional studies examining connected speech are therefore needed to allow a more accurate evaluation of language abilities.

Another limitation of the available literature is the heterogeneity of tests used (Table 2). Expressive and receptive vocabulary scales are likely to measure the same language construct across tests and languages (English, French, Italian) as they usually require children to name pictures or point to them. Yet this assumption cannot be made for tests measuring sentence comprehension and production. Indeed, some tests focused on syntax, morphology, or a combination of those. Sentence structures are also highly language-specific. We were able to identify deficits relative to normative data, but heterogeneity of assessments makes it difficult to isolate specific subcomponents of language that are affected in this patient population. Although we marked duplicate data when indicated by the authors, it was not always possible to identify those duplicates. Finally, we did not examine language with respect to IQ or verbal memory, which may correlate with some deficits.[6]

In group studies, score ranges and variation were not always provided which restricted our ability to estimate performance distribution.[28] Additionally, some of our mean z-scores for certain aetiologies were calculated on fewer than five participants, increasing the risk of bias and limiting generalization.

There was a high proportion of high and medium quality papers in our review, but the lowest scoring categories were 'patient selection process' and 'inclusion/exclusion criteria', increasing the risk of bias. The other was the 'presence of a comparison group', making it difficult to compare left hemispherectomy patients with either right-sided cases or other types of unilateral pathologies associated with intractable epilepsy. Consequently, we would recommend further studies include consecutive patients to avoid the risk of bias towards "positive recovery". Future studies should also use quantitative norm-referenced language measures to allow the reader to compare language performance to that of age matched (and ideally IQ matched) peers.

Conclusion

This large scale 30 year systematic review, compiling data from over 200 patients, regroups a large body of literature since the last review on the topic was published in 1988.[59] It also provides much needed evidence to guide clinical decisions for hemispherectomy candidates given side and aetiology. Evidence-based prognosis contributes to surgical decision for neurosurgeons and allows better counselling for families. The results will also inform the design of future prospective studies. Anticipating which language difficulties patients might experience during recovery informs speech-language therapy needs as well as support needed at school and at home.[60] Literacy development, self-esteem, quality of life and social development may be enhanced with targeted early intervention.

Appendix 1: Authors

Name	Location	Contribution
Andrea Nahum, MSci	UCL Great Ormond Street Institute of Child Health	Search, article screening, data extraction, analysis and interpretation, and quality appraisal. Drafting and revising the manuscript.
Frédérique Liégeois, PhD	UCL Great Ormond Street Institute of Child Health	Article screening, supervision of AN, revising the manuscript for intellectual content.

References

1. Hauser WA. The Prevalence and Incidence of Convulsive Disorders in Children. *Epilepsia*. 1994;35(S2):S1-S6. doi:10.1111/j.1528-1157.1994.tb05932.x
2. Spencer S, Huh L. Outcomes of epilepsy surgery in adults and children. *Lancet Neurol*. 2008;7(6):525-537. doi:10.1016/S1474-4422(08)70109-1
3. Moosa ANV, Gupta A, Jehi L, et al. Longitudinal seizure outcome and prognostic predictors after hemispherectomy in 170 children. *Neurology*. 2013;80(3). doi:10.1212/WNL.0b013e31827dead9
4. Cao K, Liu M, Wang C, et al. Five-Year Long-Term Prognosis of Epileptic Children After Hemispheric Surgery: A Systematic Review and Meta-analysis. *Medicine (Baltimore)*. 2016;95(23). doi:10.1097/MD.0000000000003743

5. Jonas R, Nguyen S, Hu B, et al. Cerebral hemispherectomy: hospital course, seizure, developmental, language, and motor outcomes. *Neurology*. 2004;62(10):1712-1721.
6. Liégeois F, Cross JH, Polkey C, Harkness W, Vargha-Khadem F. Language after hemispherectomy in childhood: Contributions from memory and intelligence. *Neuropsychologia*. 2008;46(13):3101-3107. doi:10.1016/j.neuropsychologia.2008.07.001
7. Pulsifer MB, Brandt J, Salorio CF, Vining EPG, Carson BS, Freeman JM. The Cognitive Outcome of Hemispherectomy in 71 Children. *Epilepsia*. 2004;45(3):243-254. doi:10.1111/j.0013-9580.2004.15303.x
8. Moosa AN V, Jehi L, Marashly A, et al. Long-term functional outcomes and their predictors after hemispherectomy in 115 children. *Epilepsia*. 2013;54(10):1771-1779. doi:10.1111/epi.12342
9. Mariotti P, Iuvone L, Torrioli MG, Silveri MC. Linguistic and non-linguistic abilities in a patient with early left hemispherectomy. *Neuropsychologia*. 1998;36(12):1303-1312.
10. Grosmaitre C, Jambaqué I, Dorfmuller G, et al. Exceptional verbal intelligence after hemispherotomy in a child with Rasmussen encephalitis. *Neurocase*. 2015;21(2):144-147. doi:10.1080/13554794.2013.878724
11. Anderson V, Spencer-Smith M, Wood A. Do children really recover better? Neurobehavioural plasticity after early brain insult. *Brain*. 2011;134(8):2197-2221. doi:10.1093/brain/awr103
12. Ballantyne AO, Spilkin AM, Hesselink J, Trauner DA. Plasticity in the developing brain: Intellectual, language and academic functions in children with ischaemic perinatal stroke. *Brain*. 2008;131(11):2975-2985. doi:10.1093/brain/awn176
13. Hertz-Pannier L, Chiron C, Jambaqué I, et al. Late plasticity for language in a child's non-dominant hemisphere: a pre- and post-surgery fMRI study. *Brain*. 2002;125(Pt 2):361-372. doi:10.1093/brain/awf020
14. Liégeois F, Connelly A, Baldeweg T, Vargha-Khadem F. Speaking with a single cerebral hemisphere: fMRI language organization after hemispherectomy in childhood. *Brain Lang*. 2008;106(3):195-203. doi:10.1016/j.bandl.2008.01.010
15. Higgins JP, Green S. Combining groups. In: Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration; 2011. https://handbook-5-1.cochrane.org/chapter_7/table_7_7_a_formulae_for_combining_groups.htm . Published 2011.
16. Wolraich M. *Developmental-Behavioral Pediatrics: Evidence And Practice*. 1st ed. Philadelphia: Mosby/Elsevier; 2008.
17. DistillerSR Forest Plot Generator from Evidence Partners - https://www.evidencepartners.com/resources/forest-plot-generator/
18. Liberati A, Altman D, Tetzlaff J et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. PLoS Med. 2009;6(7):e1000100. doi:10.1371/journal.pmed.1000100
19. Hrabok M, Dykeman J, Sherman EMS, Wiebe S. An evidence-based checklist to assess neuropsychological outcomes of epilepsy surgery: How good is the evidence? *Epilepsy Behav*. 2013;29(3):443-448. doi:10.1016/j.yebeh.2013.08.035
20. de Bode S, Smets L, Mathern GW, Dubinsky S. Complex syntax in the isolated right hemisphere: Receptive grammatical abilities after cerebral hemispherectomy. *Epilepsy Behav*. 2015;51:33-39. doi:10.1016/j.yebeh.2015.06.024
21. Ibrahim GM, Morgan BR, Smith M Lou, et al. Thalamocortical connectivity is enhanced following functional hemispherotomy for intractable lateralized epilepsy. *Epilepsy Behav*. 2015;51:281-285. doi:https://doi.org/10.1016/j.yebeh.2015.07.039

22. Ivanova A, Zaidel E, Salamon N, Bookheimer S, Uddin LQ, de Bode S. Intrinsic functional organization of putative language networks in the brain following left cerebral hemispherectomy. *Brain Struct Funct*. 2017;222(8):3795-3805. doi:10.1007/s00429-017-1434-y
23. Katzir T, Christodoulou JA, de Bode S. When left-hemisphere reading is compromised: Comparing reading ability in participants after left cerebral hemispherectomy and participants with developmental dyslexia. *Epilepsia*. 2016;57(10):1602-1609. doi:10.1111/epi.13507
24. Save-Pédebos J, Pinabiaux C, Dorfmuller G, et al. The development of pragmatic skills in children after hemispherotomy: Contribution from left and right hemispheres. *Epilepsy Behav*. 2016;55:139-145. doi:10.1016/j.yebeh.2015.12.013
25. Stark RE, Bleile K, Brandt J, Freeman J, Vining EPG. Speech-language outcomes of hemispherectomy in children and young adults. *Brain Lang*. 1995;51(3):406-421. doi:10.1006/brln.1995.1068
26. Villarejo-Ortega F, García-Fernández M, Fournier-Del Castillo C, et al. Seizure and developmental outcomes after hemispherectomy in children and adolescents with intractable epilepsy. *Child's Nerv Syst*. 2013;29(3):475-488. doi:10.1007/s00381-012-1949-8
27. Boatman D, Vining EP, Freeman J, Carson B. Auditory Processing Studied Prospectively in Two Hemidecorticectomy Patients. *J Child Neurol*. 2003;18(3):228-232. doi:10.1177/08830738030180030101
28. Bulteau C, Grosmaitre C, Save-Pédebos J, et al. Language recovery after left hemispherotomy for Rasmussen encephalitis. *Epilepsy Behav*. 2015;53:51-57. doi:10.1016/j.yebeh.2015.07.044
29. Bulteau C, Jambaqué I, Chiron C, et al. Language plasticity after hemispherotomy of the dominant hemisphere in 3 patients: Implication of non-linguistic networks. *Epilepsy Behav*. 2017;69:86-94. doi:10.1016/j.yebeh.2017.01.004
30. Danelli L, Cossu G, Berlingeri M, Bottini G, Sberna M, Paulesu E. Is a lone right hemisphere enough? Neurolinguistic architecture in a case with a very early left hemispherectomy. *Neurocase*. 2013;19(3):209-231. doi:10.1080/13554794.2011.654226
31. de Bode S, Chanturidze M, Mathern GW, Dubinsky S. Literacy after cerebral hemispherectomy: Can the isolated right hemisphere read? *Epilepsy Behav*. 2015;45:248-253. doi:10.1016/j.yebeh.2015.01.007
32. Dunn L, Dunn L, Styles B. *British Picture Vocabulary Scale*. 3rd ed. GL/Assessment; 2009.
33. Curtiss S, Yamada J. *The Curtiss-Yamada Comprehensive Language Evaluation (CYCLE)*. A Fine Line; 2004.
34. Dunn LM, Dunn DM. *Peabody Picture Vocabulary Test*. 4th ed. Minneapolis: Pearson Assessments; 2007.
35. Khomsi A, Khomsi J, Pasquet F, Parbeau-Gueno A. *BILO: Bilan Informatisé Du Langage Oral*. Paris: Pearson; 2007.
36. Kaplan E, Goodglass H, Weintraub S. *The Boston Naming Test*. Philadelphia, Pennsylvania: Lea & Febiger; 1983.
37. Martin N, Brownell R. *Expressive One-Word Picture Vocabulary Test*. Pearson; 2000.
38. Williams KT. *Expressive Vocabulary Test (EVT-2)*. 2nd ed. (Pearson Assessments, ed.).; 2007.
39. De Renzi E, Vignolo LA. The token test: A sensitive test to detect receptive disturbances in aphasics. *Brain*. 1962;85:665-678. doi:10.1093/brain/85.4.665
40. Bishop DVM. *Test for Reception of Grammar (TROG)*. London: Pearson; 2003.

41. Friederici AD, Brauer J, Lohmann G. Maturation of the language network: From inter- to intrahemispheric connectivities. *PLoS One*. 2011;6(6). doi:10.1371/journal.pone.0020726
42. Hickok G, Poeppel D. Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*. 2004;92(1-2):67-99. doi:10.1016/j.cognition.2003.10.011
43. Perani D, Saccuman MC, Scifo P, et al. Neural language networks at birth. *Proc Natl Acad Sci*. 2011;108(38):16056-16061. doi:10.1073/pnas.1102991108
44. Brauer J, Anwander A, Perani D, Friederici AD. Dorsal and ventral pathways in language development. *Brain Lang*. 2013;127(2):289-295. doi:10.1016/j.bandl.2013.03.001
45. Fournier NM, Calverley KL, Wagner JP, Poock JL, Crossley M. Impaired social cognition 30 years after hemispherectomy for intractable epilepsy: The importance of the right hemisphere in complex social functioning. *Epilepsy Behav*. 2008;12(3):460-471. doi:10.1016/j.yebeh.2007.12.009
46. Basheer SN, Connolly MB, Lautzenhiser A, Sherman EMS, Hendson G, Steinbok P. Hemispheric surgery in children with refractory epilepsy: seizure outcome, complications, and adaptive function. *Epilepsia*. 2007;48(1):133-140. doi:10.1111/j.1528-1167.2006.00909.x
47. Law J, Garrett Z, Nye C. Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database Syst Rev*. 2003. doi:10.1002/14651858.CD004110
48. Snowling MJ, Hulme C. Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *Br J Educ Psychol*. 2011;81(1):1-23. doi:10.1111/j.2044-8279.2010.02014.x
49. Lidzba K, Küpper H, Kluger G, Staudt M. The time window for successful right-hemispheric language reorganization in children. *Eur J Paediatr Neurol*. 2017;21(5):715-721. doi:10.1016/j.ejpn.2017.06.001
50. Carreño M, Wyllie E, Bingaman W, Kotagal P, Comair Y, Ruggieri P. Seizure outcome after functional hemispherectomy for malformations of cortical development. *Neurology*. 2001;57(2):331-333. doi:10.1212/WNL.57.2.331
51. Jin B, Krishnan B, Adler S, et al. Automated detection of focal cortical dysplasia type II with surface-based magnetic resonance imaging postprocessing and machine learning. *Epilepsia*. 2018;59(5):982-992. doi:10.1111/epi.14064
52. Curtiss S, De Bode S, Mathern GW. Spoken language outcomes after hemispherectomy: Factoring in Etiology. *Brain Lang*. 2001;79(3):379-396. doi:10.1006/brln.2001.2487
53. Everts R, Lidzba K, Wilke M, et al. Lateralization of cognitive functions after stroke in childhood. *Brain Inj*. 2010;24(6):859-870. doi:10.3109/02699051003724978
54. Dehaene-Lambertz G, Hertz-Pannier L, Dubois J, et al. Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc Natl Acad Sci*. 2006;103(38):14240-14245. doi:10.1073/pnas.0606302103
55. Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L. Functional neuroimaging of speech perception in infants. *Science (80-)*. 2002;298(5600):2013-2015. doi:10.1126/science.1077066
56. Boatman D, Freeman J, Vining E, et al. Language recovery after left hemispherectomy in children with late-onset seizures. *Ann Neurol*. 1999;46(4):579-586. doi:10.1002/1531-8249(199910)46:4<579::AID-ANA5>3.0.CO;2-K
57. Bulteau C, Delalande O, Fohlen M, Jalin C, Pinard J-M, Dellatolas G. Behaviour and cognition in children treated by hemispherotomy (Comportement et cognition chez l'enfant traité par hémisphérotomie). *ANAE - Approch Neuropsychol des Apprentissages chez l'Enfant*. 2002;14(3):204-206.

58. Vargha-Khadem F. Onset of speech after left hemispherectomy in a nine-year-old boy. Brain. 1997;120(1):159-182. doi:10.1093/brain/120.1.
59. Piacentini JC, Hynd GW. Language after dominant hemispherectomy: Are plasticity of function and equipotentiality viable concepts? *Clin Psychol Rev*. 1988;8(6):595-609. doi:10.1016/0272-7358(88)90083-9
60. Varadkar S, Tisdall M. Language after Hemispherotomy in Rasmussen Syndrome. *Pediatr Neurol briefs*. 2016;30(2):13. doi:10.15844/pedneurbriefs-30-2-4

Table 1 Study Characteristics

Study	N	Sex (M/F)	Aetiology	Age at seizure onset	Time from seizure onset to surgery	Age at surgery	Time from surgery to testing	Age at testing	Side of surgery (RX/LX)	Seizure freedom / control	Freedom with no AEDs
Boatman et al. 2003[27]	2	0/2	RS:2	LX: 8.0y RX: 8.0y	LX: 1.5y RX: 1.0y	LX: 9.5y RX: 9.0y	LX: 1.5y RX: 1.0y	LX: 11.0y RX:10.0y	1/1	2	2
Bulteau et al. 2015[28]	6	0/6	RS:6	4.9y ± 1.6y (2.6y – 7.2y)	1.3y ± 0.4y (0.8y – 2.0y)	6.1y ± 1.9y (3.4y – 8.4y)	5.6y ± 2.0y (3.0y – 8.3y)	11.7y ± 2.6y (8.4y – 14.6y)	0/6	6	6
Bulteau et al. 2017[29*]	12	5/7	RS:8; I:2; SW:1; CD:1	LX: 5.7y ± 4.8y (0y – 13.5y) RX: 5.4y ± 3.6y (2.0y – 11.2y)	LX: 5.9y ± 4.6y (0.4y – 13.6y) RX: 5.4y ± 4.3y (1.6y – 12.0y)	LX: 11.5y ± 5.4y (0.4y – 16.1y) RX: 10.9y ± 4.4y (4.0y – 15.1y)	LX: 17.9y ± 4.7y (0.9y – 15.0y) RX: 16.3y ± 3.6y (2.3y – 11.5y)	LX: 17.9y ± 2.9y (15.1y – 21.4y) RX: 16.3y ± 2.9y (12.5y – 19.5y)	5/7	12	-
Danelli et al. 2013[30]	1	1/0	Hemorrhagic angiocavernoma	No seizures	-	2.5y	14.5y	17.0y	0/1	1	1
De Bode et al. 2015[31]	7	1/6	I:7	3.1y ± 2.0y (0.1y – 5.0y)	3.9y ± 3.1y (0.9y – 9.5y)	7.0y ± 3.5y (1.0y – 10.0y)	6.6y ± 3.7y (1.0y – 11.0y)	13.6y ± 3.6y (10.0y – 21.0y)	0/7	6	5
De Bode, Smets et al. 2015[20]	10	5/5	I:7; RS:3	4.3y ± 5.0y (0.5y – 11.0y)	3.5y ± 3.1y (1.0y – 9.5y)	7.7y ± 5.4y (2.0y – 12.0y)	6.6y ± 3.2y (1.0y – 11.0y)	14.3y ± 5.6y (9.0y – 21.0y)	0/10	10	7
Grosmaitre et al. 2015[10]	1	0/1	RS:1	5.0y	1.9y	6.9y	4.2y	11.1y	0/1	1	1
Hertz-Panier et al. 2002[13*]	1	1/0	RS:1	5.5y	3.5y	9.0y	1.5y	10.5y	0/1	1	1
Ibrahim et al. 2015[21]	1	1/0	CD:1	11.0y	3.0y	14.0y	1.0y	16.0y	0/1	1	-
Ivanova et al. 2017[22]	4	1/3	I:3; CD:1	5.3y ± 4.1y (0.0y – 9.0y)	3.0y ± 2.7y (1.0y – 7.0y)	8.3y ± 4.3y (2.0y – 11.0y)	7.0y ± 4.4y (2.0y – 12.0y)	15.3y ± 4.3y (11.0y – 21.0y)	0/4	4	4
Katzir et al. 2016[23]	11	4/7	I:5; RS:6	4.8y ± 2.9y (0.6y – 11.0y)	2.0y ± 1.2y (0.6y – 4.5y)	6.8y ± 3.4y (1.2y – 12.8y)	8.5y ± 4.1y (2.1y – 14.8y)	15.3y ± 4.6y (8.9y – 22.5y)	0/11	11	7
Liegeois et al. 2008[6]	30	21/9	SW:3; RS:8; CD:1; I:12; HME:1 Hydrocephalus:1; Pachygria:1 Cyst:1; HEMIMG:1; Unknown:1 (pre/perinatal:19; postnatal:11)	LX: 3.1y ± 3.2y (0.1y – 10y) RX: 3.6y ± 2.8y (0.1y – 8.0y)	LX: 6.8y RX: 6.0y	LX: 9.8y ± 3.5y (3.2y – 15.6y) RX: 9.6y ± 4.8y (0.3y – 15.5y)	LX: 5.8y ± 4.1y (1.0y – 14.4y) RX: 5.6y± 3.3y (0.9y – 11.8y)	LX: 15.5y ± 4.0y (9.9y – 24.3y) RX: 15.2y ± 4.8y (7.9y – 21.1y)	13/17	30	-

Liegeois, Connelly et al. 2008[14]	6	5/1	I:2; RS:2; HME:1; Cyst:1	LX: 3.2y ± 3.0y (0.1y – 6.0y) RX: 3.7y ± 2.8y (0.6y – 6.0y)	LX: 5.4y ± 4.6y (0.6y – 9.8y) RX: 5.6y ± 1.9y (3.6y – 7.4y)	LX: 8.6y ± 4.1y (4.1y – 11.9y) RX: 9.3y ± 4.4y (4.2y – 11.9y)	LX: 5.8 ± 4.7y (1.6y – 10.8y) RX: 9.5y ± 1.4y (8.3y – 11.3y)	LX: 14.4y ± 0.8y (13.5y – 14.9y) RX: 18.8y ± 5.6y (12.4y – 22.8y)	3/3	6	-
Pulsifer et al. 2004[7‡]	64	25/46	RS:37; CD:20; I:7	All: 3.6y ± 3.7y	All: 3.6y ± 3.9y	All: 7.2y ± 5.1y	All: 5.4y ± 5.1y	All: 13.0y ± 8.6y	35/29	46	35
Save Pedebos et al. 2016[24]	40	18/22	CD:14; SW:4; I:9; RS13	LX: 3.6y ± 3.6y (0.0 y – 12.3y) RX: 2.1y ± 2.7y (0.0y – 8.6y)	LX: 1.5y ± 1.5y (0.3y – 5.3y) RX: 4.3y ± 4.4y (0.3y – 14.3y)	LX: 4.8y ± 4.2y (0.3y – 13.9y) RX: 6.4y ± 4.6y (0.3y – 14.3y)	LX: 7.5y ± 4.0y (0.9y – 14.6y) RX: 7.4y ± 3.9y (1.3y – 14.9y)	LX: 12.2y ± 2.6y (7.7y – 15.7y) LX: 13.8y ± 2.4y (8.1y ± 16.6y)	16/24	40	40
Stark et al. 1995[25]	7	1/6	RS:7	LX: 5.4y ± 3.1y (1.5y – 9.0y) RX: 2.5y ± 1.8y (1.0y – 4.6y)	LX: 1.9y ± 1.2y (0.3y – 3.0y) RX: 2.3y ± 1.7y (0.4y – 3.7y)	LX: 7.3y ± 3.4y (4.0y – 12.0y) RX: 4.8y ± 0.9y (3.8y – 5.7y)	LX: 3.2y ± 0.9y (2.3y – 4.1y) RX: 3.3y ± 0.3y (3.0y – 3.5y)	LX: 10.5y ± 3.7y (8.1y – 15.9y) RX: 8.2y ± 0.7y (7.3y – 8.7y)	3/4	-	-
Villarejo-Ortega et al. 2013[26]	2	2/0	RS:2	LX: 5.6y RX: 8.6y	LX: 0.6y RX: 5.2y	LX: 6.2y RX: 13.8y	LX: 5.8y RX: 1.3y	LX: 12.0y RX: 15.1y	1/1	2	0

Abbreviations: N=simple size; RS=Rasmussen Syndrome; I=infarct; CD=Cortical Dysplasia; SW=Sturge-Weber Syndrome; HME=Hemimegalencephaly; LX=left hemispherectomy; RX= right hemispherectomy. Age is expressed in years (y) in the following format: mean age ± standard deviation (range) AEDs=Antiepileptic drugs; - =information not available

* Overlapping data for one patient (L3) between Hertz-Panier et al. 2002[13] and Bulteau et al. 2017[29]

‡ Sex, age and seizure outcome data are based on the full sample, n=71

Table 2 Tests used for each language skill examined

Vocabulary		Sentence	
Receptive	*Expressive*	*Comprehension*	*Production*
British Picture Vocabulary Scale (BPVS)[32] Used in 2 studies[6,14] **Curtiss Yamada Comprehensive Language Evaluation (CYCLE) Receptive Vocabulary**[33] Used in 1 study[20] **Peabody Picture Vocabulary Test (PPVT)**[34] Used in 9 studies[7,21–23,25,27,29–31] **Recognition Vocabulary*** Used in 1 study[26] **Semantic Comprehension*** Used in 1 study[13]	**Bilan Informatisé du Langage Oral (BILO) – Lexical Production**[35] Used in 1 study[28] **Bilan Informatisé du Langage Oral (BILO) – Picture Naming**[35] Used in 1 study[10] **Boston Naming Test (BNT)**[36] Used in 1 study[30] **Expressive One Word Picture Vocabulary Test (EOWPVT)**[37] Used in 3 studies[7,25,27] **Expressive Vocabulary Test (EVT)**[38] Used in 1 study[21] **Naming*** Used in 2 studies[13,26]	**Bilan Informatisé du Langage Oral (BILO) - Oral Comprehension**[35] Used in 3 studies[10,24,28] **Syntactic Comprehension*** Used in 1 study[13] **Token Test**[39] Used in 1 study[30†] **Verbal Comprehension*** Used in 1 study[26]	**Bilan Informatisé du Langage Oral (BILO) – Statement Production**[35] Used in 1 study[28] **Bilan Informatisé du Langage Oral (BILO) - Syntactic Drafting**[35] Used in 1 study[10]

*Unpublished test or test details not specified
† Also used TROG[40] but data was not extractable

Table 3 Quality appraisal scores by study based on a modified version of the EBNP[19]

Study	1	2	3	4	5	6	7	Score
Boatman et al. 2003[27]	Y	Y	Y	N	Y	Y	Y	6
Bulteau et al. 2015[28]	Y	Y	Y	Y	Y	N	N	5
Bulteau et al. 2017[29]	Y	Y	Y	Y	Y	Y	Y	7
Danelli et al. 2013[30]	Y	Y	N	N	Y	Y	Y	5
De Bode et al. 2015[31]	Y	Y	Y	Y	Y	N	Y	6
De Bode, Smets et al. 2015[20]	Y	Y	Y	Y	Y	Y	Y	7
Grosmaitre et al. 2015[10]	Y	Y	N	N	Y	N	Y	4
Hertz-Panier et al. 2002[13]	Y	Y	N	N	Y	N	Y	4
Ibrahim et al. 2015[21]	Y	Y	N	N	Y	N	N	3
Ivanova et al. 2017[22]	Y	Y	Y	Y	Y	Y	Y	7
Katzir et al. 2016[23]	Y	Y	N	N	Y	Y	Y	5
Liegeois et al. 2008[6]	Y	Y	Y	Y	Y	Y	Y	7
Liegeois, Connelly et al. 2008[14]	Y	Y	Y	Y	Y	Y	Y	7
Pulsifer et al. 2004[7]	Y	Y	Y	Y	Y	Y	Y	7
Save Pedebos et al. 2016[24]	Y	Y	Y	N	Y	Y	Y	6
Stark et al. 1995[25]	Y	Y	Y	Y	Y	Y	Y	7
Villarejo-Ortega et al. 2013[26]	Y	Y	Y	N*	Y	Y	N	6

Abbreviations: A "Y" (=yes) indicates that the study scored 1 and a 'N' (=no) indicates that the study scored 0
*Language scores only available for two patients

Modified EBNP:
(1) Were patients described socio-demographically (i.e., age, sex, education, ethnicity, employment)? (2) Were patients described clinically (i.e., age at onset/duration of epilepsy, seizure frequency, AED use, handedness/language dominance, presurgical IQ, psychiatric comorbidity)? (3) Were the inclusion/exclusion criteria specified? ('N' for case studies); (4)Was the reason for patient selection out of a larger sample specified (e.g., from consecutive patients)? ('N' for case studies); (5) Were neurosurgical and medical data described (i.e., side, aetiology, surgical procedure, extent of resection, postsurgical neurological deficits)? (6) Was there a surgical and/or nonsurgical comparison group? (left hemispherectomy vs right hemispherectomy, other surgery or patients who did not undergo surgery); (7) Was a score reported numerically? ('N' if in graph)

Figure 1 PRISMA flow diagram of study selection

Figure 2 Forest plots of language after hemispherectomy for all aetiologies combined

Figure 3 Forest plots of language after hemispherectomy for Rasmussen Syndrome

Figure 4 Forest plots of language after hemispherectomy for Vascular Aetiologies and Cortical Dysplasia

Figure 5 Summary of language z-scores after hemispherectomy

Figure 1 PRISMA flow diagram of study selection

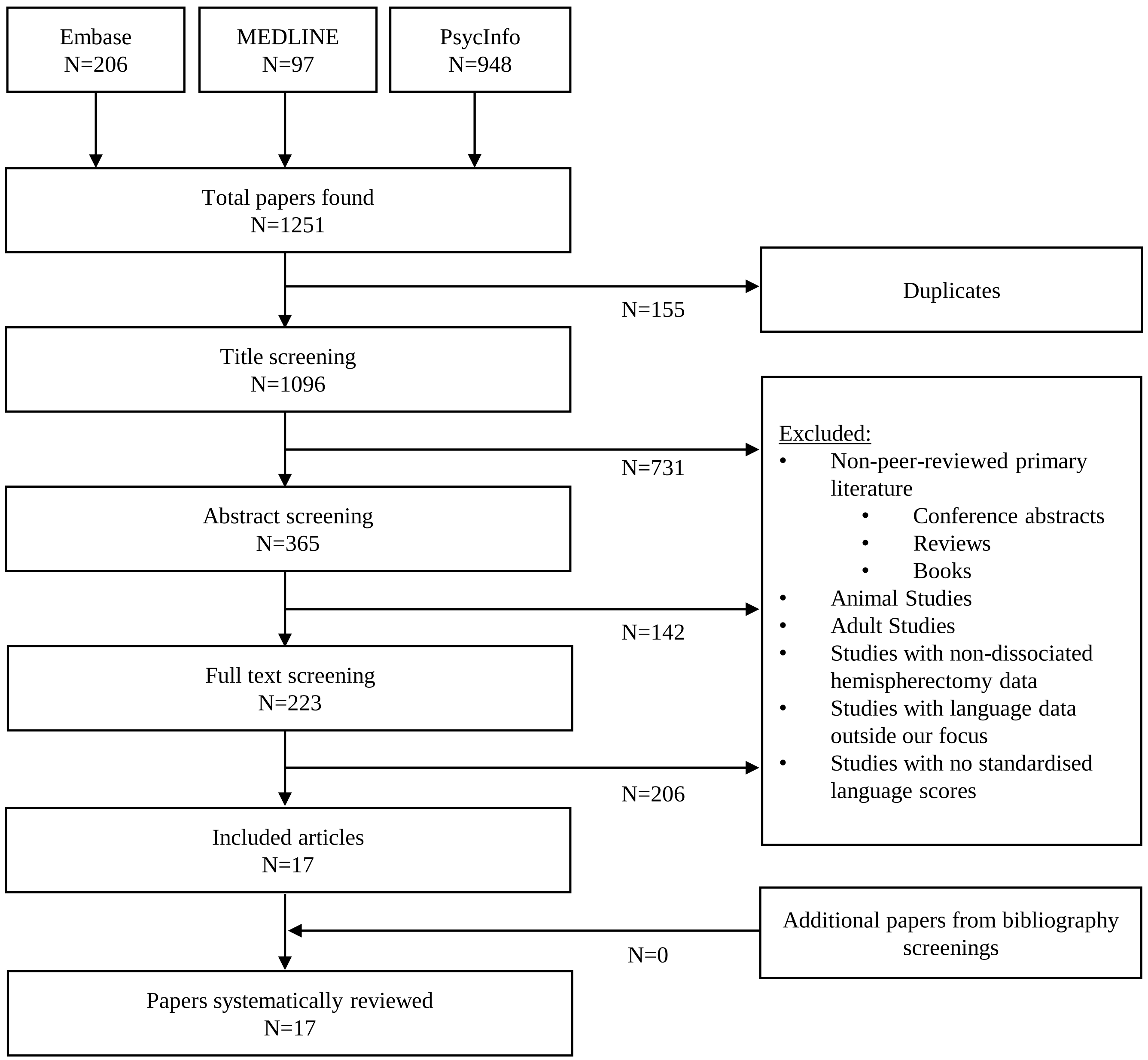
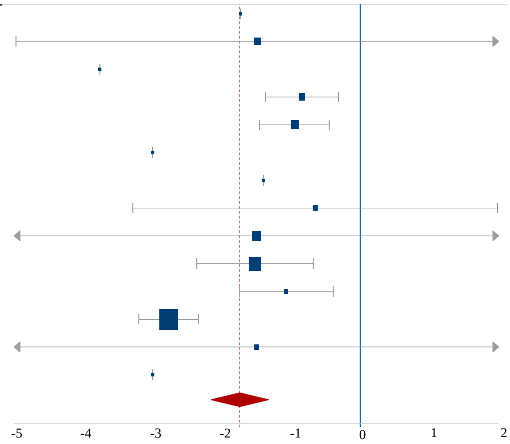


Figure 2 Forest plots of language after hemispherectomy for all aetiologies combined

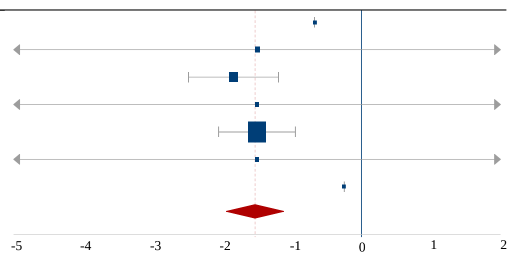
A Receptive vocabulary z-scores after left hemispherectomy

Study	N	Z score (95% CI)	Weight
Boatman et al. 2003[27]	1	-1.73	1%
Bulteau et al. 2017[29]	7	-1.12 (-4.97; 2.73)	7%
Danelli et al. 2013[30]¶	1	-3.76	1%
De Bode et al. 2015[31]	7	-0.84 (-1.37; -0.31)	7%
De Bode, Smets et al. 2015[20]	10	-0.95 (-1.45; -0.45)	10%
Hertz-Panier et al. 2003[13]	1	-3.00	1%
Ibrahim et al. 2015[21]*	1	-1.40	1%
Ivanova et al. 2017[22]	4	-0.65 (-3.28; 1.98)	4%
Katzir et al. 2016[23]	11	-1.08 (-5.07; 2.90)	12%
Liegeois et al. 2008[6]	17	-1.52 (-2.36; -0.68)	18%
Liegeois, Connelly et al. 2008[14]	3	-1.07 (-1.75 ; -0.39)	3%
Pulsifer et al. 2004[7]	29	-2.77 (-3.20; -2.34)	30%
Stark et al. 1995[25]	4	-2.30 (-8.56 ; 3.96)	4%
Villarejo-Ortega et al. 2013[26]	1	-3.00	1%
Weighted Mean	97	-1.74** (-2.02; -1.47)	100%



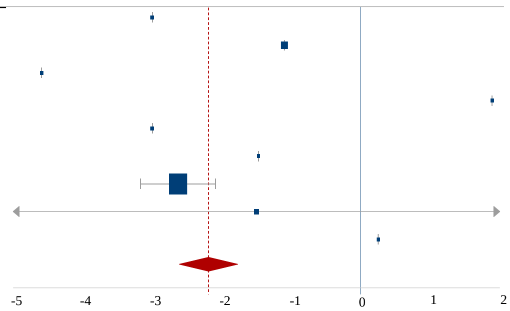
B Receptive vocabulary z-scores after right hemispherectomy

Study	N	Z score (95% CI)	Weight
Boatman et al. 2003[27]	1	-0.67	2%
Bulteau et al. 2017[29]	5	-0.59 (-6.12; 4.95)	8%
Liegeois et al. 2008[6]	13	-1.84 (-2.49; -1.19)	21%
Liegeois, Connelly et al. 2008[14]	3	-2.07 (-11.29; 7.16)	5%
Pulsifer et al. 2004[7]	35	-1.50 (-2.05; -0.95)	57%
Stark et al. 1995[25]	3	-2.33 (-13.00; 8.34)	5%
Villarejo-Ortega et al. 2013[26]*	1	-0.25	2%
Weighted Mean	61	-1.53** (-1.91; -1.16)	100%



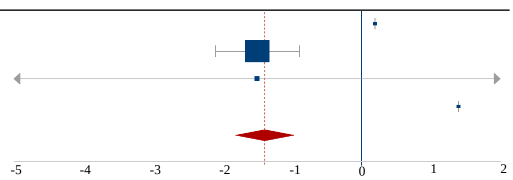
C Expressive vocabulary z-scores after left hemispherectomy

Study	N	Z score (95% CI)	Weight
Boatman et al. 2003[27]	1	-3.00	2%
Bulteau et al. 2015[28]*§	6	-1.10	15%
Danelli et al. 2013[30]¶	1	-4.59	2%
Grosmaitre et al. 2015[10]†	1	1.89	2%
Hertz-Panier et al. 2002[13]	1	-3.00	2%
Ibrahim et al. 2015[21]	1	-1.47	2%
Pulsifer et al. 2004[7]	26	-2.63 (-3.17; -2.09)	63%
Stark et al. 1995[25]	3	-1.67 (-9.11 ; 5.77)	8%
Villarejo-Ortega et al. 2013[26]*	1	0.25	2%
Weighted Mean	41	-2.19** (-2.66; -1.72)	100%



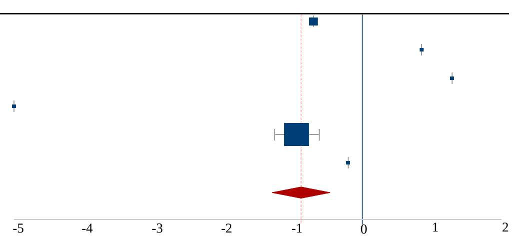
D Expressive vocabulary z-scores after right hemispherectomy

Study	N	Z score (95% CI)	Weight
Boatman et al. 2003[27]	1	0.20	3%
Pulsifer et al. 2004[7]	33	-1.49 (-2.10; -0.89)	86%
Stark et al. 1995[25]	3	-1.71 (-8.96 ; 5.54)	8%
Villarejo-Ortega et al. 2013[26]*	1	1.40	3%
Weighted Mean	38	-1.39** (-1.94; -0.83)	100%



E Sentence comprehension z-scores after left hemispherectomy

Study	N	Z score (95% CI)	Weight
Bulteau et al. 2015[28]*§	6	-0.70	18%
Danelli et al. 2013[30]‡¶	1	0.85	3%
Grosmaitre al. 2015[10]†	1	1.29	3%
Hertz-Panier et al. 2002[13]	1	-5.00	3%
Save Pedebos et al. 2016[24]	24	-0.94 (-1.26; -0.62)	70%
Villarejo-Ortega et al. 2013[26]	1	-0.20	3%
Weighted Mean	34	-0.88** (-1.25; -0.51)	100%

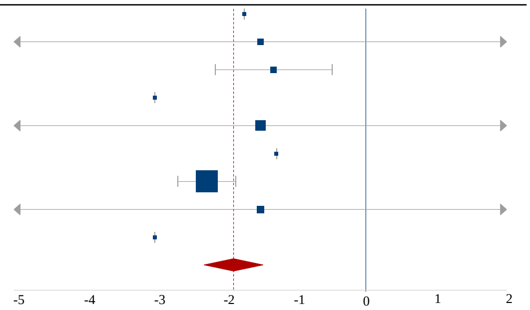


¶ Scores were calculated using control group data from the same study as population test norms were unavailable
*Mean was roughly estimated from a graph as data were not provided in text format
§ No estimate of variation was provided in this study, therefore confidence intervals (CI) could not be calculated
† Clinical history suggests this child was right dominant for language functions at birth
‡ Non-standardised scores, calculated using mean and standard deviation of the control group. Patient evaluated earlier than for other language measures at 8.3y
** Mean is significantly lower than for the general population, $p<0.001$

Figure 3 Forest plots of language after hemispherectomy for Rasmussen Syndrome

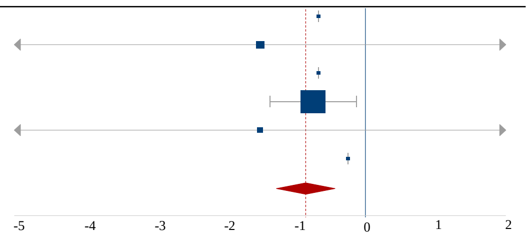
A Receptive vocabulary z-scores after left hemispherectomy

Study	N	Z score (95% CI)	Weight
Boatman et al. 2003[27]	1	-1.73	3%
Bulteau et al. 2017[29]	3	-0.69 (-7.53; 6.15)	8%
De Bode, Smets et al. 2015[20]	3	-1.31 (-2.14; -0.48)	8%
Hertz-Panier et al. 2002[13]	1	-3.00	3%
Katzir et al. 2016[23]	6	-1.23 (-7.27; 4.80)	17%
Liegeois, Connelly et al. 2008[14]	1	-1.27	3%
Pulsifer et al. 2004[7]	16	-2.26 (-2.67; -1.85)	44%
Stark et al. 1995[25]	4	-2.30 (-8.56 ; 3.96)	11%
Villarejo-Ortega et al. 2013[26]*	1	-3.00	3%
Weighted Mean	36	-1.88** (-2.22; -1.54)	100%



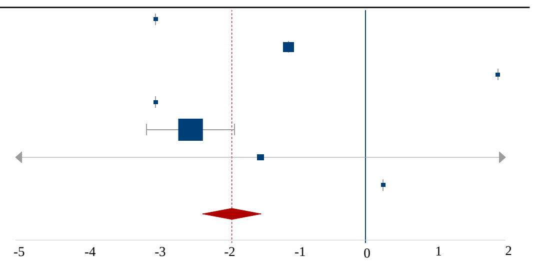
B Receptive vocabulary z-scores after right hemispherectomy

Study	N	Z score (95% CI)	Weight
Boatman et al. 2003[27]	1	-0.67	3%
Bulteau et al. 2017[29]	5	-0.59 (-6.12; 4.95)	16%
Liegeois, Connelly et al. 2008[14]	1	-0.67	3%
Pulsifer et al. 2004[7]	21	-0.74 (-1.36; -0.13)	66%
Stark et al. 1995[25]	3	-2.33 (-13.00; 8.34)	9%
Villarejo-Ortega et al. 2013[26]*	1	-0.25	3%
Weighted Mean	32	-0.85** (-1.32; -0.38)	100%



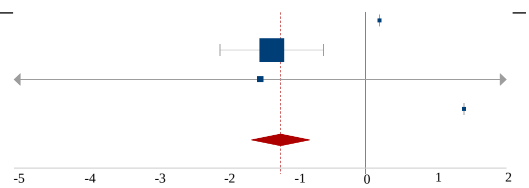
C Expressive vocabulary z-scores after left hemispherectomy

Study	N	Z score (95% CI)	Weight
Boatman et al. 2003[27]	1	3.00-	4%
Bulteau et al. 2015[28]*§	6	-1.10	20%
Grosmaitre et al. 2015[10]†	1	1.89	4%
Hertz-Panier et al. 2002[13]	1	-3.00	4%
Pulsifer et al. 2004[7]	16	-2.50 (-3.13; -1.87)	55%
Stark et al. 1995[25]	3	-1.67 (-9.11; 5.77)	10%
Villarejo-Ortega et al. 2013[26]*	1	0.25	4%
Weighted Mean	29	-1.91** (-2.44; -1.38)	100%



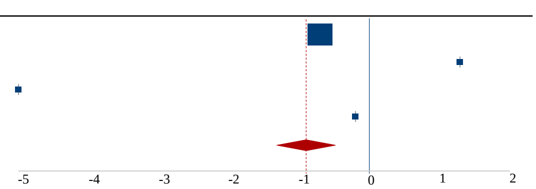
D Expressive vocabulary z-scores after right hemispherectomy

Study	N	Z score (95% CI)	Weight
Boatman et al. 2003[27]	1	0.20	4%
Pulsifer et al. 2004[7]	20	-1.33 (-2.07; -0.60)	80%
Stark et al. 1995[25]	3	-1.71 (-8.96; 5.54)	12%
Villarejo-Ortega et al. 2013[26]*	1	1.40	4%
Weighted Mean	25	-1.21** (-1.86; -0.55)	100%



E Sentence comprehension z-scores after left hemispherectomy

Study	N	Z score (95% CI)	Weight
Bulteau et al. 2015[28]*§	6	-0.70	67%
Grosmaitre al. 2015[10]†	1	1.29	11%
Hertz-Panier et al. 2002[13]	1	-5.00	11%
Villarejo-Ortega et al. 2013[26]*	1	-0.20	11%
Weighted Mean	9	-0.90 (-1.99; 0.19)	100%



*Mean was roughly estimated from a graph as data were not provided in text format

§ No estimate of variation was provided in this study, therefore confidence intervals (CI) could not be calculated

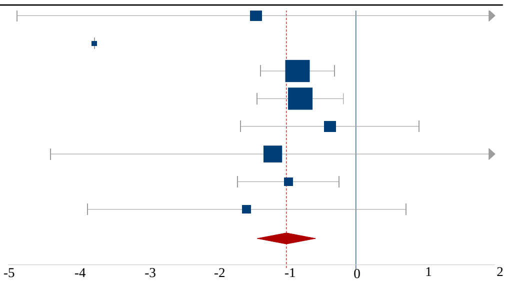
† Clinical history suggests this child was right dominant for language functions at birth

** Mean is significantly lower than for the general population , $p<0.001$

Figure 4 Forest plots of language after hemispherectomy for Vascular Aetiologies and Cortical Dysplasia

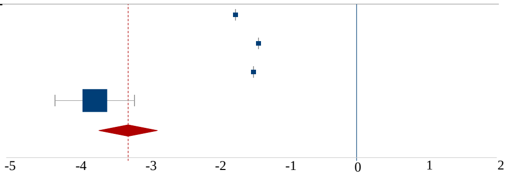
A Receptive vocabulary z-scores after left hemispherectomy for Vascular Aetiologies

Study	N	Z score (95% CI)	Weight
Bulteau et al. 2017[29]	3	-1.36 (-4.87; 2.16)	10%
Danelli et al. 2013[30]¶	1	-3.76	3%
De Bode et al. 2015[31]	7	-0.84 (-1.37; -0.31)	23%
De Bode, Smets et al. 2015[20]	7	-0.80 (-1.42; -0.18)	23%
Ivanova et al. 2017[22]	3	-0.38 (-1.66; 0.91)	10%
Katzir et al. 2016[23]	5	-0.91 (-4.39; 2.58)	17%
Liegeois, Connelly et al. 2008[14]	2	-0.97 (-1.70; -0.24)	7%
Pulsifer et al. 2004[7]	2	-1.57 (-3.86; 0.72)	7%
Weighted Mean	30	-1.00** (-1.33; -0.68)	100%



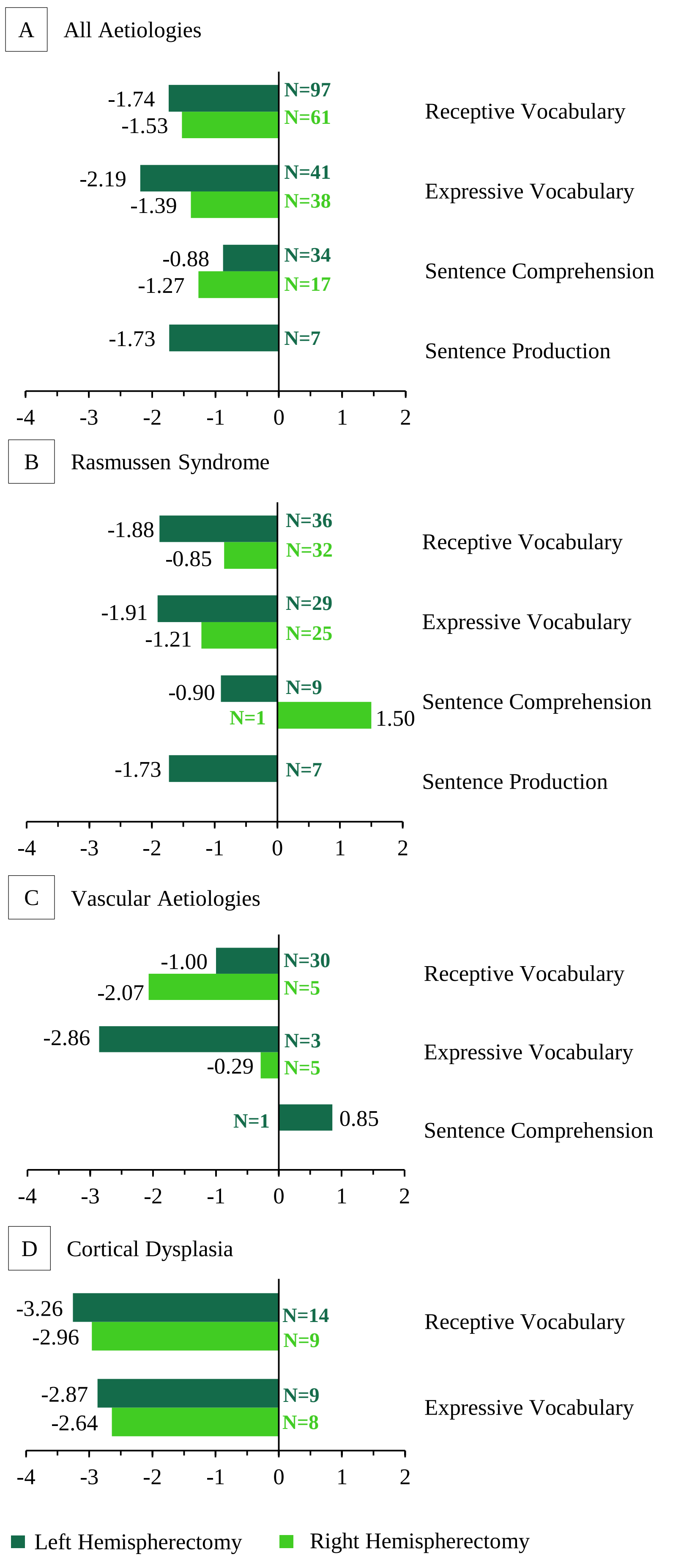
B Receptive vocabulary z-scores after left hemispherectomy for Cortical Dysplasia

Study	N	Z score (95% CI)	Weight
Bulteau et al. 2017[29]	1	-1.73	7%
Ibrahim et al. 2015[21]*	1	-1.40	7%
Ivanova et al. 2017[22]	1	-1.47	7%
Pulsifer et al. 2004[7]	11	-3.73 (-4.30; -3.17)	79%
Weighted Mean	14	-3.26** (-3.92; -2.61)	100%



¶ Scores were calculated using control group data from the same study as population test norms were unavailable
*Mean was roughly estimated from a graph as data were not provided in text format
** Mean is significantly lower than for the general population, $p<0.001$

Figure 5 Summary of language z-scores after hemispherectomy



N indicates sample size